

# **The National R&D Evaluation System in the UK**

## **- Meta-evaluation and Applicability to the Korean Case -**

A thesis submitted to the University of Manchester for the degree of  
PhD in the Faculty of Social Sciences and Law

2003

**Chan-goo Yi**

**Policy Research in Engineering, Science and Technology  
(PREST)**

ProQuest Number: 10756478

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10756478

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code  
Microform Edition © ProQuest LLC.

ProQuest LLC.  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 – 1346

✕

1424360

✓

SEP 11 1998  
1424360

# TABLE OF CONTENTS

<b>TABLE OF CONTENTS</b>	<b>2</b>
<b>LIST OF FIGURES</b>	<b>7</b>
<b>LIST OF TABLES</b>	<b>8</b>
<b>ABSTRACT</b>	<b>12</b>
<b>DECLARATION AND COPYRIGHT</b>	<b>13</b>
<b>ACKNOWLEDGEMENT</b>	<b>14</b>
<b>LIST OF ABBREVIATIONS</b>	<b>15</b>
<b>THE AUTHOR</b>	<b>19</b>
<b>Chapter 1. Introduction</b>	<b>20</b>
1.1. Research Purposes	20
1.2. Research Necessities	22
1.3. Scope and Level of Research Object	24
1.4. Research Contents and Methods	29
1.5. Structure of the Thesis	32
<b>Chapter 2. Theoretical Background</b>	<b>34</b>
2.1. Introduction	34
2.2. Policy Evaluation	34
2.2.1. Definition of Policy Evaluation	35
2.2.2. Types of Policy Evaluation	37
2.3. R&D Evaluation	39
2.3.1. Definition of R&D Evaluation	40
2.3.2. Types of R&D Evaluation	40
2.3.3. Major Components of R&D Evaluation	41
2.4. Relations between Policy Evaluation and R&D Evaluation	45



<b>Chapter 3. Research Framework : Meta-evaluation Model</b>	<b>48</b>
3.1. Introduction	48
3.2. Review of Existing Research	49
3.3. Definition of Meta-evaluation	63
3.3.1. Review of the Existing Definition of Meta-evaluation	63
3.3.2. New Definition of Meta-evaluation	65
3.4. Components of Meta-evaluation	66
3.4.1. Limitations of Existing Research	66
3.4.2. Discussion of New Components in Meta-evaluation	67
3.4.2.1. Evaluation Paradigm	67
3.4.2.2. Evaluation Resources	69
3.4.2.3. Evaluation Performance	71
3.4.2.4. Evaluation Utilisation	73
3.5. Design of Research Framework : Meta-evaluation Model	75
 <b>Chapter 4. Case Study of the UK National R&amp;D Evaluation System</b>	 <b>77</b>
4.1. Introduction	77
4.2. Current State of the R&D Evaluation Systems in the UK	77
4.2.1. Government Organisations for SET	78
4.2.2. Brief History of the Development of R&D Evaluation	90
4.2.3. Current Hierarchy of National R&D Evaluation Systems	97
4.2.4. Guidance and Consultation Documents for R&D Evaluation	107
4.3. Recent Evaluation Exercises in S&T and R&D Sectors	119
4.3.1. Evaluation of Science and Research Policy : Cross-cutting Review of Science and Research by HM Treasury (2000)	120
4.3.2. Evaluation of Research Councils : Quinquennial Review of the Six Grant Awarding Research Councils by the DTI and the OST (2001)	127
4.3.3. Evaluation of Research Institutes : The Institute Assessment Exercise by the BBSRC (2001)	138
4.3.4. Evaluation of S&T Programme : Evaluation of Funding for UK Civil Space Activity by the DTI and other Departments (2001)	147

4.3.5. Evaluation of R&D Programme : Review of the Gastrointestinal Response to Food (GRF) Programme by the BBSRC (2000) -----	161
4.3.6. Synthesis and Conclusion -----	171
<b>Chapter 5. Case Study of the Korean National R&amp;D Evaluation System ---</b>	<b>173</b>
5.1. Introduction -----	173
5.2. Current State of the R&D Evaluation Systems in Korea -----	174
5.2.1. Government Organisations for S&T -----	174
5.2.2. Brief History of the Development of R&D Evaluation -----	190
5.2.3. Current Hierarchy of National R&D Evaluation Systems -----	197
5.2.4. Current Referential Statutes for R&D Evaluation -----	209
5.3. Recent Evaluation Exercises in S&T and R&D Sectors -----	223
5.3.1. Evaluation of S&T and R&D Policy : Evaluation of Major Policy Issues of the MOST by the OPM/the OGPC (2001) -----	224
5.3.2. Evaluation of R&D Programme : Evaluation of the National R&D Programmes by the NSTC (2001) -----	235
5.3.3. Evaluation of Research Councils : Evaluation of Five Research Councils by the OGPC (2001) -----	246
5.3.4. Evaluation of Government-supported Research Institutes (GRIs) : Evaluation of Member Research Institutes by the KOCI (2002) -----	254
5.3.5. Evaluation of R&D Projects : Evaluation of the Creative Research Initiatives (CRI) by the MOST (2000) -----	263
5.3.6. Synthesis and Conclusion -----	275
<b>Chapter 6. Meta-evaluation and Comparison of the UK and the Korean Cases ---</b>	<b>277</b>
6.1. Introduction -----	277
6.2. Meta-evaluation of the UK R&D Evaluation System -----	278
6.2.1. Meta-evaluation of the Cross-cutting Review of Science and Research in 2000 -----	279

6.2.2. Meta-evaluation of the Quinquennial Review of the Six Grant Awarding Research Councils in 2001 .....	286
6.2.3. Meta-evaluation of the Institute Assessment Exercise (IAE) by the BBSRC in 2001 .....	291
6.2.4. Meta-evaluation of the Evaluation of Funding for UK Civil Space Activity by the DTI in 2001 .....	297
6.2.5. Meta-evaluation of the Review of the Gastrointestinal Response to Food (GRF) Programme by the BBSRC in 2001 .....	303
6.2.6. Section Conclusion .....	308
6.3. Meta-evaluation of the Korean R&D Evaluation System .....	311
6.3.1. Meta-evaluation of the Evaluation of Major Policy Issues of the MOST in 2001 .....	311
6.3.2. Meta-evaluation of the Evaluation of National R&D Programmes in 2001 .....	320
6.3.3. Meta-evaluation of the Evaluation of Five Research Councils in 2001 ---	327
6.3.4. Meta-evaluation of the Evaluation of Member Research Institutes by the KOCI in 2002 .....	333
6.3.5. Meta-evaluation of the Evaluation of the Creative Research Initiatives (CRI) by the MOST in 2000 .....	339
6.3.6. Section Conclusion .....	345
6.4. Comparison of the UK and the Korean R&D Evaluation Systems .....	349
6.4.1. Comparison of the Structure of National R&D Evaluation System .....	350
6.4.2. Comparison of the Evaluation System for S&T and R&D Policy .....	355
6.4.3. Comparison of the Evaluation System for S&T or R&D Programme .....	360
6.4.4. Comparison of the Evaluation System for Research Councils .....	364
6.4.5. Comparison of the Evaluation System for Research Institutes under the Research Councils .....	369
6.4.6. Comparison of the Evaluation System for R&D Projects .....	374
6.4.7. Synthesis .....	379

<b>Chapter 7. Suggestions for Applicability of the UK Case to the Korean Case</b>	<b>386</b>
7.1. Introduction	386
7.2. Suggestion for Applicability of the UK System to the Korean System	386
7.2.1. Discussion of the Structure of National R&D Evaluation System	387
7.2.2. Discussion of the Evaluation System for S&T and R&D Policy	389
7.2.3. Discussion of the Evaluation System for S&T or R&D Programme	391
7.2.4. Discussion of the Evaluation System for the Research Councils	393
7.2.5. Discussion of the Evaluation System for the Research Institutes	395
7.2.6. Discussion of the Evaluation System for the R&D Projects	397
7.3. Chapter Conclusions	398
 <b>Chapter 8. Conclusions</b>	 <b>399</b>
8.1. Synthesis of Research	399
8.1.1. Summary of the Thesis	399
8.1.2. Key Suggestions for the Applicability of Aspects of the UK Case to the Korean Case	407
8.2. Limitations and Further Research	408
8.3. Contributions	410
8.3.1. Academic Perspective : Development of Research in the Area of Meta-evaluation and R&D Evaluation	410
8.3.2. Practical Perspective : Development of the R&D Evaluation Systems of the UK and Korea	413
 <b>Bibliography</b>	 <b>415</b>
 <b>List of Interviewees</b>	 <b>433</b>

# LIST OF FIGURES

(Figure 1-1) Hierarchy of the UK R&D Evaluation System and its Research Scope -----	26
(Figure 1-2) Hierarchy of the Korean R&D Evaluation System and its Research Scope ---	28
(Figure 1-3) Research Flow -----	33
 (Figure 2-1) Main Areas of Policy Evaluation -----	 36
 (Figure 4-1) Overview of the UK System of Governance of SET -----	 80
(Figure 4-2) Overview of the Dual Support System of HEIs -----	88
(Figure 4-3) Hierarchy of National R&D Evaluation System in the UK -----	98
 (Figure 5-1) Overview of the Korean System of Governance of S&T -----	 176
(Figure 5-2) Hierarchy of National R&D Evaluation System in Korea -----	199
(Figure 5-3) Objects of the Programme Evaluation and the Project Review -----	238

# LIST OF TABLES

<Table 1-1> Research Scope and Level in the UK Case Study -----	29
<Table 1-2> Research Scope and Level in the Korean Case Study -----	29
<Table 2-1> Types of Policy Evaluation and Main Evaluation Object -----	39
<Table 2-2> Types of R&D Evaluation and its Principal Purposes -----	41
<Table 2-3> Areas of Policy Evaluation and its Main Types -----	46
<Table 2-4> Areas of R&D Evaluation and its Main Types -----	46
<Table 3-1> Literature Regarding Meta-evaluation -----	49
<Table 3-2> Types of Case Studies in Meta-evaluation -----	50
<Table 3-3> Cook and Gruder's Meta-evaluation Models -----	52
<Table 3-4> Larson and Berliner's Evaluation Components -----	56
<Table 3-5> The Joint Committee's Evaluation Standards -----	59
<Table 3-6> Research Framework : Meta-evaluation Model -----	75
<Table 4-1> Typology of Government Organisations for SET in the UK -----	81
<Table 4-2> Number of UK Research Councils Institutes -----	90
<Table 4-3> Development of Policy Evaluation and R&D Evaluation in the UK -----	96
<Table 4-4> R&D Evaluation Systems in the Civil Departments in the UK -----	106
<Table 4-5> Guidance and Consultation Documents for R&D Evaluation -----	108
<Table 4-6> Recent Evaluation Exercises in S&T and R&D Sectors (UK) -----	120
<Table 4-7> Breakdown of the Working Group Members in Quinquennial Review ---	130
<Table 4-8> Organisations of Submission in Stage One of Quinquennial Review -----	133
<Table 4-9> Breakdown of the Respondents in Stage Two of Quinquennial Review --	133
<Table 4-10> Evaluation Object of the IAE 2001 -----	139
<Table 4-11> Information Requirements in the IAE 2001 -----	142
<Table 4-12> Evaluation Criteria and Indicators in the IAE 2001 -----	145
<Table 4-13> Breakdown of the Manpower Commitment in the AU -----	154

<Table 4-14> Membership of the Agri-Food Committee in the BBSRC .....	165
<Table 4-15> Detailed Information from the Questionnaire for GRF Evaluation .....	167
<Table 4-16> Evaluation Criteria and Indicator for GRF Programme .....	170
<Table 4-17> Synthesis of the Evaluation Exercises in S&T and R&D Sector of the UK .....	172
<Table 5-1> Typology of Government Organisations for S&T in Korea .....	178
<Table 5-2> Current State of the GRIs in Korea (2001) .....	187
<Table 5-3> Development of Policy Evaluation and R&D Evaluation in Korea .....	196
<Table 5-4> Current Referential Statutes for R&D Evaluation in Korea .....	210
<Table 5-5> Recent Evaluation Exercises in S&T and R&D Sectors (Korea) .....	224
<Table 5-6> Members and their Organisations of the PEC and the Support Team .....	230
<Table 5-7> Evaluation Criteria and Items for Major Policy Issues in 2001 Exercise ..	232
<Table 5-8> Membership of the Evaluation Committee for sub R&D Programmes .....	242
<Table 5-9> Membership of the Review Committee for R&D Projects .....	242
<Table 5-10> Recommendations for Improvement of R&D Programmes .....	245
<Table 5-11> Recommendations for Improvement of R&D Projects .....	246
<Table 5-12> Main and Sub Evaluation Objects for Research Councils .....	248
<Table 5-13> Membership of the Evaluation Committee for Research Councils .....	251
<Table 5-14> Evaluation Criteria and Indicators for Research Councils .....	253
<Table 5-15> Main and Sub Evaluation Objects for GRIs under the KOCI .....	256
<Table 5-16> Major Information Requirements for the 2002 Evaluation Exercise .....	259
<Table 5-17> Membership of Evaluation Committee in the KOCI's 2002 Exercise ---	260
<Table 5-18> Evaluation Criteria and Indicators for the GRIs under the KOCI .....	262
<Table 5-19> Project Numbers and Fund Investment for the CRI (1997 to 2001) .....	265
<Table 5-20> Research Areas and Performers for 1997 Projects under the CRI .....	266
<Table 5-21> Detailed Membership of the PMC for the CRI .....	269
<Table 5-22> Information Requirements for the 2000 Exercise for the CRI .....	270
<Table 5-23> The Role of the Evaluator Groups in the 2000 Exercise of the CRI .....	271
<Table 5-24> Membership of Evaluation Panel for the CRI 2000 Exercise .....	271
<Table 5-25> Details of International Evaluators in 2000 CRI Exercise .....	272

<Table 5-26> Evaluation Criteria and Indicators in 2000 CRI Exercise .....	273
<Table 5-27> Evaluation Results in the 2000 CRI Exercise .....	274
<Table 5-28> Synthesis of the Evaluation Exercises in S&T and R&D Sector of Korea .....	276
<Table 6-1> Current Types of R&D Evaluation Exercises in the UK and Korea .....	351
<Table 6-2> Relations among Sub Evaluation Exercises in the UK and Korea .....	352
<Table 6-3> Evaluation Timing of Sub Evaluation Exercises in the UK and Korea ....	354
<Table 6-4> Comparison of Evaluation Paradigm for S&T and R&D Policy in the UK and Korea .....	356
<Table 6-5> Comparison of Evaluation Resources for S&T and R&D Policy in the UK and Korea .....	357
<Table 6-6> Comparison of Evaluation Performance for S&T and R&D Policy in the UK and Korea .....	358
<Table 6-7> Comparison of Evaluation Utilisation for S&T and R&D Policy in the UK and Korea .....	359
<Table 6-8> Comparison of Evaluation Paradigm for S&T or R&D Programme in the UK and Korea .....	360
<Table 6-9> Comparison of Evaluation Resources for S&T or R&D Programme in the UK and Korea .....	362
<Table 6-10> Comparison of Evaluation Performance for S&T or R&D Programme in the UK and Korea .....	363
<Table 6-11> Comparison of Evaluation Utilisation for S&T or R&D Programme in the UK and Korea .....	364
<Table 6-12> Comparison of Evaluation Paradigm for Research Councils in the UK and Korea .....	365
<Table 6-13> Comparison of Evaluation Resources for Research Councils in the UK and Korea .....	366
<Table 6-14> Comparison of Evaluation Performance for Research Councils in the UK and Korea .....	368



<Table 6-15> Comparison of Evaluation Utilisation for Research Councils in the UK and Korea -----	369
<Table 6-16> Comparison of Evaluation Paradigm for Research Institutes in the UK and Korea -----	370
<Table 6-17> Comparison of Evaluation Resources for Research Institutes in the UK and Korea -----	371
<Table6-18> Comparison of Evaluation Performance for Research Institutes in the UK and Korea -----	373
<Table6-19> Comparison of Evaluation Utilisation for Research Institutes in the UK and Korea -----	374
<Table6-20> Comparison of Evaluation Paradigm for Research Projects in the UK and Korea -----	375
<Table6-21> Comparison of Evaluation Resources for Research Projects in the UK and Korea -----	376
<Table6-22> Comparison of Evaluation Performance for Research Projects in the UK and Korea -----	377
<Table6-23> Comparison of Evaluation Utilisation for Research Projects in the UK and Korea -----	378
<Table6-24> Comprehensive Comparison of Structure of R&D Evaluation Systems in the UK and Korea -----	379
<Table6-25> Comprehensive Comparison of Evaluation System for S&T and R&D Policy -----	379
<Table6-26> Comprehensive Comparison of Evaluation System for S&T or R&D Programme -----	380
<Table6-27> Comprehensive Comparison of Evaluation System for Research Councils -----	380
<Table6-28> Comprehensive Comparison of Evaluation System for Research Institutes -----	381
<Table6-29> Comprehensive Comparison of Evaluation System for R&D Projects -----	381

## ABSTRACT

The major concern of this research is to suggest development for the Korean national R&D evaluation system, taking the UK case as a benchmarking model. In order to do this, two major research purposes have been established. The first is to investigate the strengths and limitations of the UK and the Korean R&D evaluation systems. The second is to discuss the extent to which the UK exercises could be applied to the Korean cases.

For this purposes, I firstly conducted the two case studies of the UK and the Korean national R&D evaluation system each of which covered five evaluation exercises : (1) the evaluation of S&T and R&D policy, (2) the evaluation of S&T or R&D programmes, (3) the evaluation of research councils, (4) the evaluation of research institutes under the research councils and (5) the evaluation of R&D projects. Then, in order to identify the strengths and limitations of the UK and the Korean systems as well as the differences between the two countries in relation to the five evaluation exercises, the meta-evaluations and the comparisons of those meta-evaluations in the two countries were conducted.

From the above meta-evaluations and comparisons, it is evident that there is some possibility of applying some elements of the UK evaluation systems and its exercises to the Korean cases. First of all, as a precondition for the applicability, I suggest that the four Korean R&D evaluation systems, excluding that for R&D projects, should extend the evaluation timing, depending on the attributes of the individual R&D evaluation systems. Next, it appears that, among the five evaluation exercises in the UK, the evaluation system of the S&T programme can provide many useful references for the future development of the Korean system. There is also a possibility of applying aspects of the UK exercise to the Korean case for both the research councils and the research institutes. Lastly, among 11 evaluation components, both the evaluation objects and the utilisation type of the evaluation findings of the UK exercises are expected to be the best references for the Korean cases.

## **DECLARATION**

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university.

## **COPYRIGHT**

Copyright in text of this thesis rests with the Author. Copies (by any process) either in full, or of extracts, may be made only in accordance with instructions given by the Author and lodged in the John Rylands University Library of Manchester. Details may be obtained from the Librarian. This page must form part of any such copies made. Further copies (by any process) of copies made in accordance with such instructions may not be made without the permission (in writing) of the Author.

The ownership of any intellectual property rights which may be described in this thesis is vested in the University of Manchester, subject to any prior agreement to the contrary, and may not be made available for use by third parties without the written permission of the University, which will prescribe the terms and conditions of any such agreement.

Further information on the conditions under which disclosures and exploitation may take place is available from the Head of the Department of PREST (Policy Research in Engineering, Science and Technology).

## ACKNOWLEDGEMENT

First and foremost, I would like to express great gratitude to my supervisor, Professor Luke Georghiou and to my advisor, Ms Kate Barker. Without their excellent guidance and advice, this thesis would not have been successfully completed. In extending my sincere thanks to them, I wish to say that I shall remember their constant support and favours during my research in the University. I also owe many things to all faculty members of the PREST, in particular, Dr. Maria Nedeva, Dr. Pual Cummingham, Dr. Kieron Flanagan and Dr. Lawrence Green. Through discussions with them, I was able to get a lot of information essential for the development of my thesis.

In addition, I have found much support from outside the PREST. Firstly, I have to thank Professor Keun-bok Kang, my first supervisor in Korea, who has encouraged me throughout the whole period of undertaking my second degree in the UK. I would also like to thank the British Council in Seoul, which gave me a good opportunity to do research at the University of Manchester.

Finally, special thanks are due to my wife, Dr. Pal-nam Kim, who has continuously supported and encouraged me, and to my two lovely daughters, Sally (Gyeong-eun) and Jenny (Seok-eun), who have done the best for their lives in the UK, even though they have had less assistance from me than I would have liked to give. During those busy and stressful times in Manchester, I have once again realised how valuable and important my family is in my life.

## LIST OF ABBREVIATIONS

AATSR : Advanced Along Track Scanning Radiometer  
ACO : Administrative Co-ordination Office, The Republic of Korea  
AEA : American Evaluation Association  
AO : Assessment Office  
APU : Assessment and Policy Unit  
AU : Assessment Unit  
BAI : Board of Audit and Inspection, The Republic of Korea  
BBSRC : Biotechnology and Biological Sciences Research Council, The UK  
BNSC : British National Space Centre, The UK  
CCLRC : Council for the Central Laboratory of the Research Councils, The UK  
CERCs : Chief Executives of the Research Councils  
CF : Competitive Funding  
CSA : Chief Scientific Adviser, The UK  
CSG : Core Strategic Grants  
CSR : Comprehensive Spending Review  
CVCP : Committee of Vice Chancellors and Principals  
DEFRA : Department for Environment, Food and Rural Affairs, The UK  
DENI : Department of Education for Northern Ireland, The UK  
DfES : Department for Education and Skills, The UK  
DfT : Department for Transport, The UK  
DGRC : Director General of the Research Councils  
DoH : Department of Health, The UK  
DTI : Department of Trade and Industry, The UK  
ELDO : European Launcher Development Organisation  
EPB : Economic Planning Board, The Republic of Korea  
EPIC : Evaluation and Policy Improvement Committee, The UK  
EPSRC : Engineering and Physical Sciences Research Council, The UK  
ERC : Engineering Research Centre, The Republic of Korea

ESA : European Space Agency  
 ESRC : Economic and Social Research Council, The UK  
 ESRO : European Space Research Organisation  
 ESPRIT : European Strategic Programme for R&D in IT  
 FMI : Financial Management Initiative  
 FRID : Finance Regulation and Industry Directorate, The UK  
 GIIs : Government-invested Institutions  
 GRIs : Government-supported Research Institutes  
 HEFCs : Higher Education Funding Councils, The UK  
 HEFCE : Higher Education Funding Council for England, The UK  
 HEIs : Higher Education Institutions, The UK  
 HPEB : Health Technology Planning and Evaluation Board, The Republic of Korea  
 HRI : Horticulture Research International, The UK  
 HSE : Health and Safety Executive, The UK  
 IAE : Institute Assessment Exercise  
 IITA : Institute of Information Technology Assessment, The Republic of Korea  
 IT : Information Technology  
 ITEP : Korea Institute of Industrial Technology Evaluation and Planning  
 KAIST : Korea Advanced Institute of Science and Technology  
 KCESRI : Korea Councils of Economic and Social Research Institutes  
 KCHSRI : Korea Councils of Humanities and Social Research Institutes  
 KIST : Korea Institute of Science and Technology  
 KISTEP : Korea Institute of Science and Technology Evaluation and Planning  
 K-JIST : Kwang-ju Institute of Science and Technology, The Republic of Korea  
 KOCI : Korea Research Council for Industrial Science and Technology  
 KORP : Korea Research Council of Public Science and Technology  
 KOSEF : Korea Science and Engineering Foundation  
 KPS : Korean Physical Society  
 KRCF : Korea Research Council of Fundamental Science and Technology  
 KRF : Korea Research Foundation  
 KSEA : Korean-American Scientist and Engineers Association

KT : Knowledge Transfer

MAF : Ministry of Agriculture and Forestry, The Republic of Korea

MIC : Ministry of Information and Communication, The Republic of Korea

MOCIE : Ministry of Commerce, Industry and Energy, The Republic of Korea

MOCT : Ministry of Construction and Transportation, The Republic of Korea

MoD : Ministry of Defence, The UK

MOE : Ministry of Education and Human Resources Development, The Republic of Korea

MOER : Ministry of Energy and Resources, The Republic of Korea

MOHW : Ministry of Health and Welfare, The Republic of Korea

MOST : Ministry of Science and Technology, The Republic of Korea

MPB : Ministry of Planning and Budget, The Republic of Korea

MRC : Medical Research Council, The UK

NAO : National Audit Office, The UK

NDPBs : Non-Departmental Public Bodies

NERC : Natural Environment Research Council, The UK

NIHEC : Northern Ireland Higher Education Council, The UK

NRIs : National Research Institutes, The Republic of Korea

NSTC : National Science and Technology Council, The Republic of Korea

OGPC : Office for Government Policy Co-ordination, The Republic of Korea

OPIs : Output and Performance Indicators

OPM : Office of the Prime Minister, The Republic of Korea

OST : Office of Science and Technology, The UK

PACST : Presidential Advisory Councils on Science and Technology, The Republic of Korea

PAR : Programme Analysis and Review

PC : Programme Coordinator

PEC : Policy Evaluation Committee, The Republic of Korea

PIs : Principal Investigators

PMC : Planning and Management Committee

PPARC : Particle Physics and Astronomy Research Council, The UK

PREST : Policy Research in Engineering, Science and Technology, The UK

PSAs : Public Service Agreements  
 PSD : Public Service Directorate  
 PSREs : Public Sector Research Establishments, The UK  
 RAE : Research Assessment Exercise  
 R&D : Research and Development  
 RCs : Research Councils  
 ROAME : Rationale, Objective, Appraisal, Monitoring and Evaluation  
 ROAMEF : Rationale, Objectives, Appraisal, Monitoring, Evaluation and Feedback  
 RRC : Regional Research Centre, The Republic of Korea  
 QR : Quality-related Research  
 QS : Quality of Science  
 SABRIs : Scottish Agricultural and Biological Research Institutes, The UK  
 S&T : Science and Technology  
 SCI : Science Citation Index  
 SDAs : Service Delivery Agreements  
 SEB : Science and Engineering Base  
 SEBG : Science and Engineering Base Group  
 SEERAD : Scottish Executive Environment and Rural Affairs Department, The UK  
 SERC : Science and Engineering Research Council, The UK  
 SESC : Space Evaluation Steering Committee  
 SET : Science, Engineering and Technology  
 SEWG : Space Evaluation Working Group  
 SME : Small and Medium-sized Enterprise  
 SPRU : Science Policy Research Unit, The UK  
 SR : Spending Review  
 SRC : Science Research Centre, The Republic Korea  
 STAU : Science and Technology Assessment Unit, The UK  
 STEPI : Science and Technology Policy Institute, The Republic Korea  
 TESE : Technology, Economics, Statistics and Evaluation Directorate, The UK  
 TDSTG : Trans-departmental Science and Technology Group  
 VG : Visiting Group



## THE AUTHOR

Chan-goo Yi received his first PhD degree, which dealt with the topic of the meta-evaluation of the Korean national R&D programme in the area of information and communication, in 1997 from the Department of Public Administration in the Chungnam National University, Daejeon, The Republic of Korea. His major research topics are policy evaluation, R&D evaluation and S&T policy. His research papers, in particular, relating to the Korean R&D evaluation systems and their exercises, have been also published in an international journal such as *Research Evaluation* as well as in some Korean leading journals, including *Korean Policy Studies Review*, *Korean Public Administration Quarterly* and *Korean Journal of Policy Analysis and Evaluation*.

Meanwhile, he has been working as a principal researcher in ETRI (Electronics and Telecommunications Research Institute), one of the Korean government supported research institutes, since 1988. Just before joining the PREST in September 2000, he was a team manager of the Project Evaluation Team responsible for the evaluation of research units, research projects and researcher's performance in the ETRI.

# CHAPTER 1. INTRODUCTION

## 1. 1. Research Purposes

Since the late 1980s, the UK government has conceived of policy evaluation as a useful policy means for the reform and improvement of performance in the public sector and has made a great effort to institutionalise policy evaluation at the level of central government (HM Treasury, 1988 ; 1997). In the context of policy evaluation, from the late 1980s, the UK government has also started to evaluate public research and development (R&D) programmes or projects and those research institutes that have been major performers in such programmes or projects (Cabinet Office, 1989). As a result, it is argued that, at the moment, the UK has the most developed national R&D evaluation system of all the European countries, and that this system has had direct or indirect influences on some developing countries, including Korea, as well as on other developed countries (Hills and Dale, 1995 : 35 ; Lee et al., 1994a). Also, the UK has played an important role in the evaluation of the EU's large-scale R&D programmes, such as the Framework Programme, EUREKA<sup>1</sup> and the European Strategic Programme for R&D in IT (ESPRIT) (Georghiou and Cameron, 1987 ; Georghiou and Meyer-Kramer, 1992 ; Dale and Barker, 1994 ; Georghiou, 1995).

In Korea, until the late 1990s, the Ministry of Science and Technology (MOST), the Ministry of Commerce, Industry and Energy (MOCIE) and the Ministry of Information and Communication (MIC) had evaluated their own major national R&D programmes or projects. However, each of the above-mentioned government ministries has evaluated its major national R&D programmes, not from the perspective of R&D policies and science and technology (S&T) policies nationwide, but from its own point of view. Furthermore, they have conducted their evaluations, not at the level of the programmes themselves, but at

---

<sup>1</sup> EUREKA is not an acronym but its name was originally related to the concept of a European Research Co-ordination Agency.

that of projects. Meanwhile, regarding the evaluation of research institutes, the MOST had conducted evaluations of only its affiliated Government-supported Research Institutes (GRIs) between 1991 and 1997. However, as there was no legal requirement for such evaluations, which were thus performed only due to the MOST's internal needs, the evaluation of research institutes, for example, the GRIs and the National Research Institutes (NRIs), was not the general case in Korea.

For the above-mentioned reasons, R&D evaluation activities in Korea have not achieved their purposes, which are, for example, the effective allocation of R&D resources and the production of useful information for the improvement of R&D activity. In order to deal with this deficiency, the Korean government revised the "Special Act for Scientific and Technological Innovation" in February 1998.<sup>2</sup> On the basis of this law, the Minister of the MOST, as the secretariat of the National Science and Technology Council (NSTC), was actually authorised to evaluate all national R&D programmes (MOST, 1998). Also, in January 1999, the Korean government enacted the "Act on the Establishment, Operation and Promotion of the GRIs", and gave three research councils<sup>3</sup> the missions of evaluating the performance of R&D projects and the management systems of affiliated GRIs (OPM, 2001c). In addition to these laws, in January 2001, the Korean government enacted a new law known as the "Framework Act on the Government Performance Evaluation" (OPM, 2001a), to improve the policy evaluation system and its implementation including R&D evaluation activity at the national level.

Meanwhile, the Korean government has intended to utilise the evaluation findings as the basic information for both the development of national R&D programmes or projects and the improvement of performance in public sector research institutes. In accordance with this government policy, in Korea, there are huge demands for the rational and effective establishment and implementation of a national R&D evaluation system, from scientists and researchers in GRIs as well as from policy makers in government.

---

<sup>2</sup> This law was succeeded to a new law known as the "Framework Act on Science and Technology" in 2001.

<sup>3</sup> Three research councils are : (1) The Korea Research Council of Fundamental Science and Technology (KRCF), (2) The Korea Research Council of Public Science and Technology (KORP) and (3) The Korea Research Council for Industrial Science and Technology (KOI).

Hence, the major concern of this research is to suggest developments for the Korean national R&D evaluation system, taking the UK case as a benchmarking model. In order to do this, two major research purposes have been established. The first is to investigate the strengths and limitations of the UK and the Korean R&D evaluation systems and their exercises. The second is to discuss the extent to which the UK's theories and experiences in R&D evaluation could be applied to the Korean case, and if there is any possibility of applying aspects of the UK system to the Korean system, to suggest the preconditions for the successful adjustment of the UK system to the Korean situation.

## **1. 2. Research Necessities**

Regarding R&D evaluation, there are diverse types of research, including theories of R&D evaluation, case studies relating to individual countries, comparative studies between two or more countries, proposals for improvement of R&D evaluation systems. As a major concern is the fact that the national R&D evaluation systems of the UK and Korea are different from each other due to the differing political systems and S&T policies of the two countries, I shall only review literature which has dealt with this matter. Through a review of the literature of this type, I can uncover the research topics which have been missed in the existing research and also suggest both the necessity for and the importance of studying this topic.

In the UK, since the early 1980s, there has been a great accumulation of research on this subject by the Policy Research in Engineering, Science and Technology (PREST) at the University of Manchester and, in particular in early stage, the Science Policy Research Unit (SPRU) at the University of Sussex (Hills, 1995). On the basis of the research results of these two universities, the UK has been able to play a major role in the evaluation of the EU's R&D programmes, as I have already indicated.

However, despite the great quantity of existing research, there is only a little in relation to complete R&D evaluation systems at the national level (Hills and Dale, 1995 ; Barker, 1997 ; Barker and Lloyd, 1997 ; Hills, 2001) or in relation to individual R&D evaluation

systems at government department level (Cunion, 1995 ; Barber, 1999a ; Barber, 2001). This research has been focused on some parts of R&D evaluation systems, such as the development of R&D evaluation, the method of R&D evaluation, the utilisation of R&D evaluation and so on, rather than on national R&D evaluation systems as a whole. Hence, it is not easy to acquire a comprehensive and systematic understanding of the UK R&D evaluation system by using these. As a result, such research materials as exist are neither objective nor sufficient for a foreign country to be able to use in order to take the UK R&D evaluation system as a benchmarking model. In particular, if some developing countries, including Korea, are to adopt the R&D evaluation system of another country having more advanced systems and experiences, such as the UK, extreme caution and much preparation are necessary. For this reason, I think that researchers or policy makers of foreign countries will use the results of this research as a useful basis for understanding the UK R&D evaluation system more precisely.

Meanwhile, in Korea, there has also been a lot of research on this subject by researchers in both universities and GRIs. This research can be categorised as follows : general studies on the R&D evaluation systems (Hwang et al., 2000 ; Lee, 2000), evaluations of major national R&D programmes by the ministry responsible for them (MOCIE, 1992 ; MOST, 1997 ; MIC et al., 1998), evaluations of individual national R&D projects (Lee et al., 1993 ; Hong et al., 1998), proposals for improvement of the evaluation activity in individual R&D programmes (Kim et al., 1992 ; Park, 1995 ; Lee et al., 1994b ; Lee et al., 1995 ; Lee et al., 1996 ; Yi and Kang, 1998 ; Kim, 1999), proposals for future development of the evaluation of the institutes in the R&D sector (Rho et al., 1996 ; Yi and Kang, 2000b). This existing research has normally introduced R&D evaluation systems in the developed countries such as the UK, the USA and Japan. The findings of this research have also suggested improvement in R&D evaluation activity in individual R&D programmes or projects and future development for R&D evaluation systems as a whole. Therefore, even in its relatively short period of carrying out R&D evaluation, by using the above-mentioned existing research, Korea has established, developed and implemented a rational and effective R&D evaluation system relatively quickly.

However, even though existing research has contributed considerably to the development of the Korean R&D evaluation system, some of this research has merely proposed the transfer of systems and components of the system from one country to another (i.e. Korea), without taking into account their strengths and weaknesses, or considering the historical and cultural background of each country and the implications this may have for the necessary adoption of such systems when they are so transferred. Taking these factors into account can avert in advance the failure of imported systems and the introduction into developing countries of problems that have already been experienced in developed countries. However, even so, incompatibilities will still occur between detailed components of R&D evaluation system as they are being implemented.

Therefore, in this research, I shall first identify the strengths and limitations of the national R&D evaluation system in the UK and Korea. Next, on the basis of this, I shall also investigate how the UK R&D evaluation system could be applied to the Korean system, even though the background to the latter is different. Of course, there may be some difficulties in adopting a single country as the case study for a benchmarking. To overcome this deficiency, I shall point out the weaknesses and anticipated problems of the policy alternatives recommended. Also, through this work, I shall try to minimise the errors that could happen when Korea adopts the UK case.

### **1. 3. Scope and Level of Research Object**

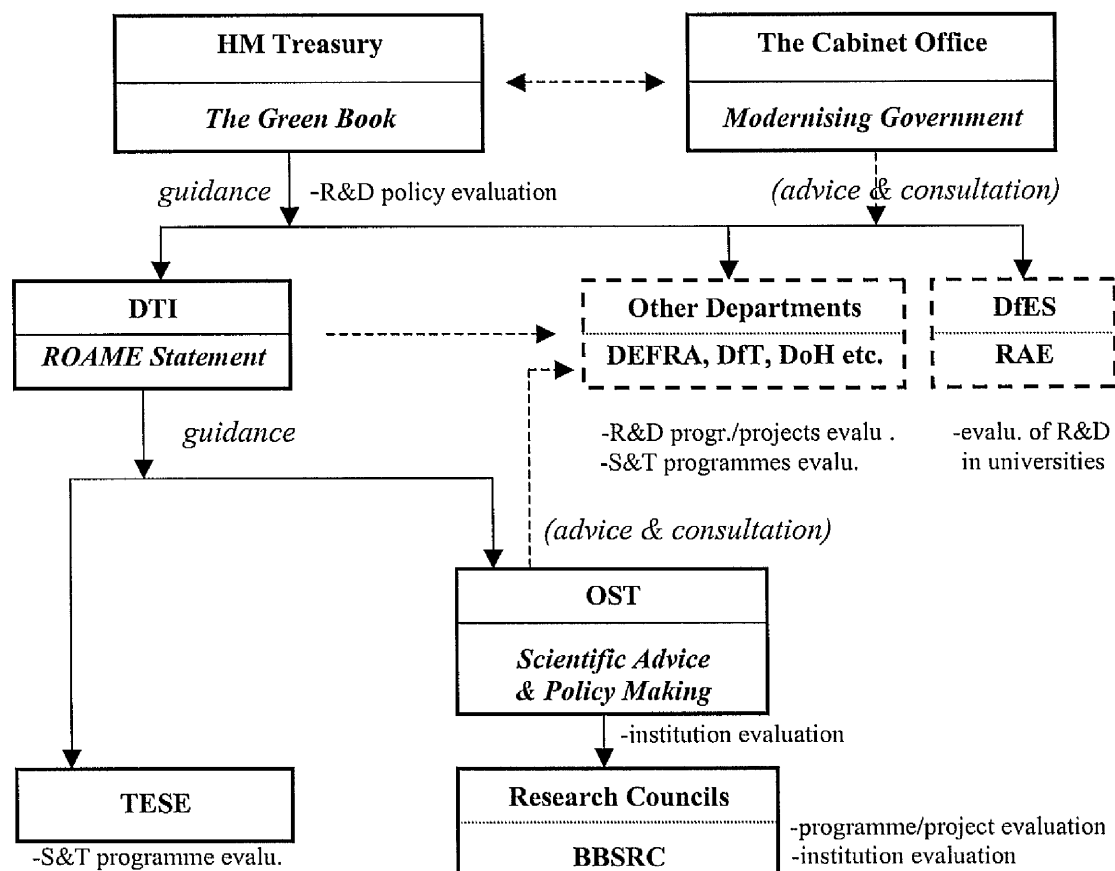
As this research deals with the R&D evaluation system in both the UK and Korea at the national level, it will not cover R&D evaluation activity in the private sector. However, as there are diverse types of R&D evaluation at the national level in both countries, it is not in fact possible even to cover all R&D evaluation activities in the public sector. Thus, I shall confine the scope and level of the research object as indicated in the following.

First, I shall discuss the research scope which will decide the extent of the research object. For a general understanding of the R&D evaluation system in the UK, it is firstly necessary to describe the UK R&D activities in the public sector. It is normally argued that, in the UK,

government departments which are responsible for performing public sector R&D activities are divided into two types (Hills and Dale, 1995 : 36-37 ; Cunningham, 1998b : 70, 74). One type of government department is represented by the Department of Trade and Industry (DTI) and the Department for Education and Skills (DfES), which have carried out basic and strategic R&D programmes for future national competitiveness and economic development as a whole. The R&D programmes that the DTI and the DfES are in charge of have normally been implemented by public research institutes and universities. Under the DTI, there is the Office of Science and Technology (OST), which has missions to co-ordinate all national R&D programmes among the departments concerned, and to obtain the science budget and distribute it to the research councils (RCs) and other science-related organisations. Meanwhile, the seven research councils, having been given their science budget by the OST, have supported the R&D programmes of public sector research institutes and universities. Of these, the Biotechnology and Biological Sciences Research Council (BBSRC) is one of the largest and the leading councils in the UK, and is also responsible for supporting the R&D programmes which are considered as crucial areas in the UK SET sector (Cunningham, 1998c : 179). Another type of government department is represented by the Department for Environment, Food and Rural Affairs (DEFRA), the Department for Transport (DfT), the Ministry of Defence (MoD) and the Department of Health (DoH) which have carried out R&D programmes or projects for their own policy purposes

In the UK, there are some current guidelines on how the central government is to evaluate the diverse types of national R&D programmes or projects (Hills, 2001). Of these, *Appraisal and Evaluation in Central Government : The Green Book* (HM Treasury, 1997), *The Modernising Government* (Cabinet Office, 1999) and *Guidelines 2000 : Scientific Advice and Policy Making* (OST, 2000b) are the most important ones. According to these guidelines, individual government departments have to set up their own guidelines for the assessment and evaluation of the R&D programmes and projects under their control. Though there has been a lot of guidance from some departments, it is normally argued that *ROAME (Rational, Objective, Appraisal, Monitoring and Evaluation) Statement* of the DTI (2000a) has been the most systematic guidance and has had the most influence on other

departments (Hills and Dale, 1995 ; Hills, 2001). Meanwhile, the evaluation of R&D activities in universities supported by the DfES has been conducted every five years in the form of the Research Assessment Exercise (RAE). The RAE has inherent within itself a logic and methodology which are a little different from those of general R&D evaluation approaches (Cunningham, 1998c). As a result, in the UK, HM Treasury, the Cabinet Office, the DTI, the OST and the Research Councils have played a major role in evaluating R&D programmes, projects and the research institutions in the public sector. Thus, while carrying out the case study on the UK national R&D evaluation system, I shall confine the research scope to the organisations outlined with a bold line in (Figure 1-1). These include HM Treasury, the Cabinet Office, the DTI/TESE (Technology, Economics, Statistics and Evaluation) and the Research Councils (in particular, the BBSRC).



(Figure 1-1) Hierarchy of the UK R&D Evaluation System and its Research Scope

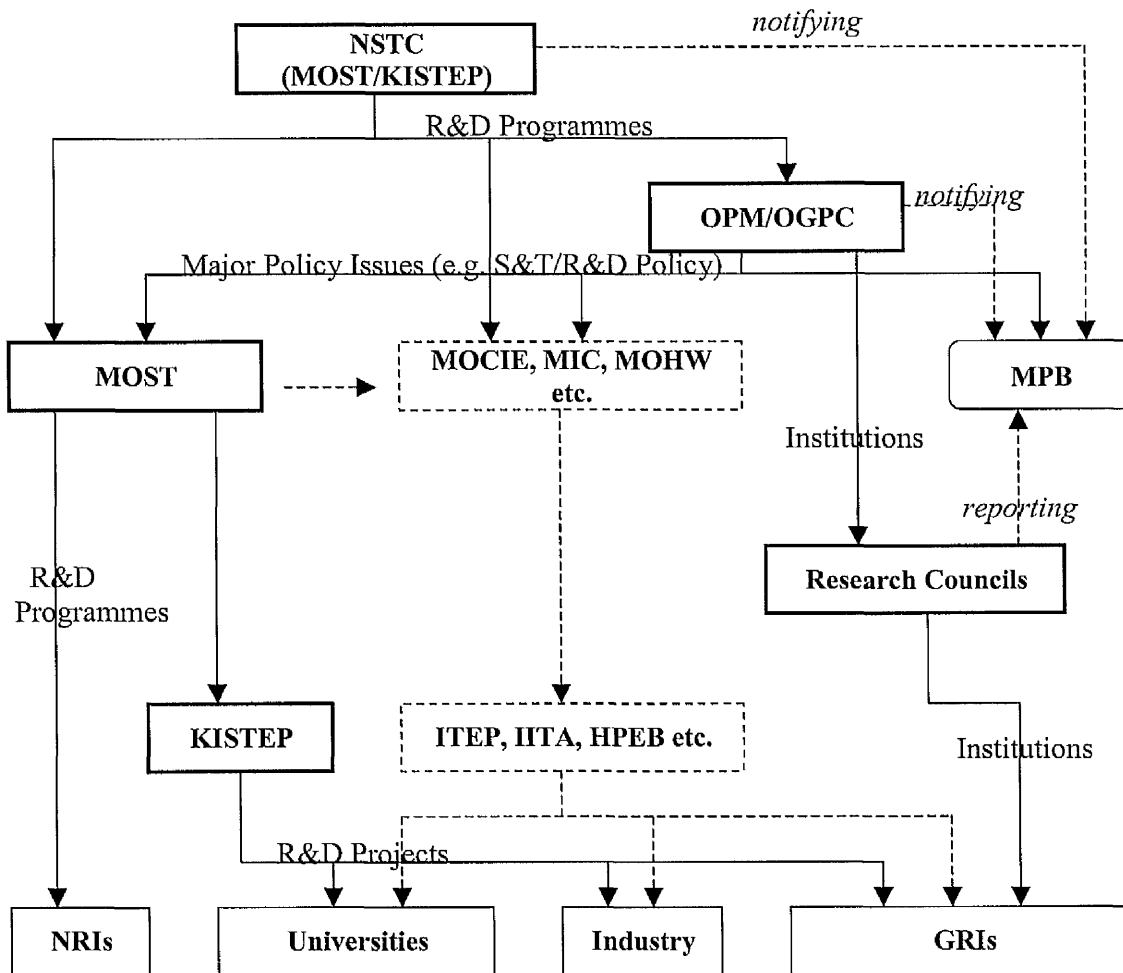


Second, I shall discuss the research level, which will determine the upper and the lower limits of the research object. In general, the objects of R&D evaluation can be classified into R&D policy, R&D programmes/projects, research institutions and researchers (Gibbons and Georghiou, 1987 ; Cabinet Office, 1989). According to this classification, the major objects of this research will be the evaluation system of S&T and R&D policy, S&T and R&D programmes, research institutions and R&D projects. However, even though there are a lot of evaluations of researchers' performance, this aspect will be excluded from the research object, because the basic concept and methods of evaluating a researcher are different from those of evaluating the other four evaluation objects mentioned above (Cabinet Office, 1989). As a result, the upper research level will be the evaluation of S&T and R&D policy and the lower one will be the evaluation of R&D projects.

For the case study on the Korean national R&D evaluation system, I shall adopt the same scope and level of research objects as in the case of the UK. As one of the main contents of this research is a comparative study between the UK and Korea in the area of R&D evaluation, it is necessary to set up the scope and level of research in such a way that they can be compared to each other directly or indirectly.

In the Korean case study, the research scope will be restricted to organisations outlined with a bold line in (Figure 1-2). These include the NSTC, the MOST, the Office of the Prime Minister (OPM)/the Office for Government Policy Co-ordination (OGPC), the Research Councils (in particular, the KOCI) and R&D Management Institutes (in particular, the KISTEP : the Korea Institute of Science and Technology Evaluation and Planning). In Korea, in addition to these, there are some ministries, such as the MOCIE and the MIC, which are responsible for the national R&D programmes and projects. Under the control of these ministries, there are some R&D Management Institutes, for example, the Korea Institute of Industrial Technology Evaluation and Planning (ITEP) and the Institute of Information Technology Assessment (IITA), which have missions to select, manage and evaluate the R&D projects. However, these organisations will be excluded from the research scope, because, in Korea, the MOST and the KISTEP are more important than other ministries and organisations in the area of R&D evaluation, just as the DTI and the

BBSRC are in the UK case. Finally, in the Korean case, the research level will be restricted to the same S&T and R&D policy evaluation, R&D programme/project evaluation and research institutions evaluation as in the UK case study.



(Figure 1-2) Hierarchy of the Korean R&D Evaluation System and its Research Scope

Consequently, the scope and level of the research object in the UK and the Korean case studies can be summarised as shown in <Table 1-1> and <Table 1-2>, respectively.

**<Table 1-1> Research Scope and Level in the UK Case Study**

Level Scope	S&T and R&D Policy Evaluation	S&T and R&D Programme/Project Evaluation	Institution Evaluation
HM Treasury	■ □	□	□
Cabinet Office	□	□	□
OST	□	□	■
DTI	□	■	■
BBSRC	-	■	■

□ : advice and consultation for evaluation

■ : execution of evaluation

**<Table 1-2> Research Scope and Level in the Korean Case Study**

Level Scope	S&T and R&D Policy Evaluation	R&D Programme/ Project Evaluation	Institution Evaluation
NSTC	□	■	□
OPM/OGPC	■	-	■
MOST	□	■	■
KOCI	-	□	■
KISTEP	-	■	□

□ : advice and consultation for evaluation

■ : execution of evaluation

## 1. 4. Research Contents and Methods

It can normally be argued that the R&D evaluation system of a certain country has been developed against the background of its original political and administrative system, as well as its science and technology policy, all of which may be different, and in some cases, extremely different from those of other countries. Extreme caution and much preparation are necessary if a certain country is to adopt the R&D evaluation system of other countries which have more advanced systems and experiences. Hence, as already indicated, the major

concern of this research is to suggest developments for the Korean national R&D evaluation system, taking the UK case as a benchmarking model.

As a result, this thesis will comprise eight chapters, the contents of which will now be outlined. I shall base the research in this thesis on a literature review and interviews. The former will be mainly used for the theoretical discussion and design of the research framework, and partly used as a background to the case study in both the UK and Korea. The interviews will be used to complement what is unclear or unconfirmed in the literature regarding the national R&D evaluation systems of the two countries. Meanwhile, in analysing the research data and materials relating to the UK and the Korean case studies, I shall adopt both a qualitative research method and a quantitative one.

Chapter 1, the current chapter, outlines the purposes and need for research, the scope and level of the research object and the methods and contents of proposed research.

In Chapter 2, I shall discuss the theoretical background of this research, that is, the theory of policy evaluation and R&D evaluation, which are slightly different in their historical development as they are contained in different academic fields. I shall first proceed to the definition and types of policy evaluation. Then, I shall deal with the definition of R&D evaluation and its types and components. Finally, on the basis of the previous discussions, the relations between policy evaluation and R&D evaluation will be briefly addressed. In order to do this, I shall review the existing research articles that are mainly concerned with theoretical debates in the area of both policy evaluation and R&D evaluation.

Chapter 3 will deal with the research framework of this thesis. In order to achieve the two different research purposes already described in Section 1.1, I shall use a meta-evaluation methodology for the first one and a comparative analytical methodology for the second. Of these, the meta-evaluation will be the main research framework. A meta-evaluation is simply defined as the evaluation of an evaluation (Orata, 1940 ; Scriven, 1969). In more detail, it is the evaluation of the findings, contents and processes of a certain evaluation during and after the evaluation activities, and of the evaluation system or structure in

broader terms (Cook and Gruder, 1978 ; Stufflebeam, 1981 ; Larson and Berliner, 1983 ; Hoogerwerf, 1992 ; Rogers and Hough, 1995 ; Rogers, 1995).

There have been a lot of theoretical discussions (Cook and Gruder, 1978 ; Stufflebeam, 1981 ; Larson and Berliner, 1983 ; Chelimsky, 1987 ; Schwandt and Halpern, 1988 ; Smith and Hauser, 1990 ; Greene, 1992 ; The Joint Committee on the Standards for Educational Evaluation, 1994 ; Rogers and Hough, 1995 ; Rogers 1995) as well as case studies (Hoogerwerf, 1992 ; Georgiou, 1995 ; Yi, 1997 ; Uusikylä and Virtanen, 2000 ; Hong, 2000) in the area of meta-evaluation. However, there are few meta-evaluation studies in relation to the R&D evaluation systems themselves at the national level. Therefore, I shall review the existing research on this subject critically and comprehensively, and suggest a meta-evaluation model that comprises four main evaluative components : an evaluation paradigm, evaluation resources, evaluation performance and evaluation utilisation as a research framework, which can be usefully adopted for the analysis and judgement of R&D evaluation system at the national level.

In Chapter 4, I shall give an account of the case study on the UK national R&D evaluation system. In order to do this, firstly, I shall review the existing research articles and reports, as well as formal and informal documentation published by the government organisations concerned with R&D evaluation, such as HM Treasury, the Cabinet Office, the DTI, the OST and the research councils, including the BBSRC. Next, in order to complement what is unclear or unconfirmed in the literature review, I shall conduct interviews with previous and current staff members who have been involved in the actual evaluation exercises. Interviews will mainly be conducted face-to-face, on the telephone and by e-mail, depending on the specific interview situations.

Chapter 5 will deal with the case study on the Korean R&D evaluation system, which will be carried out using the same research methods and process as in the UK case study.

In Chapter 6, I shall use the results of the two case studies in Chapters 4 and 5 to meta-evaluate and compare the national R&D evaluation systems of the UK and Korea, using the

meta-evaluation model already designed in Chapter 3. The four main components and 11 sub components of the meta-evaluation model will be used as the criteria for meta-evaluating and comparing the cases. Through this work, I shall be able to gauge the difference between the current state of the two countries' R&D evaluation system and the ideal state of R&D evaluation system.

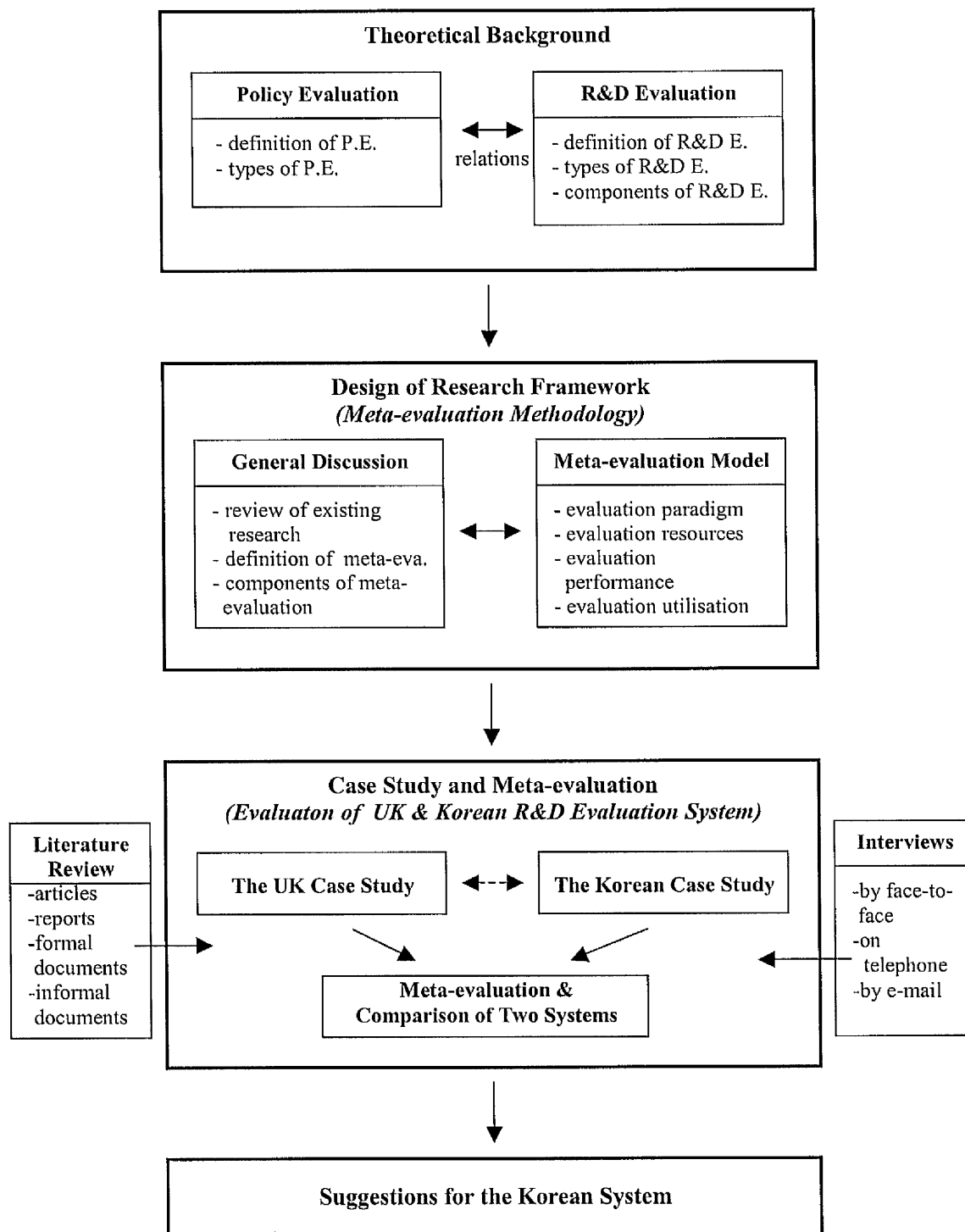
In addition, I shall attempt to extract some useful references for applying the UK exercise to the Korean case. In particular, by utilising the findings in the previous chapters, I shall first discuss the extent to which the UK exercises in R&D evaluation could be applied to the Korean case. I shall also suggest the preconditions for the UK system being successfully adjusted to the Korean situation.

Chapter 7 will suggest how the UK system can be applied in Korea. In doing this, I shall bear in mind fundamental research questions, such as whether the UK system, which is different from that of Korea in terms of its political and administrative structure, as well as in its S&T and R&D policy, could be adopted as a useful benchmarking model in Korea.

Finally, in Chapter 8, I shall firstly synthesise the overall findings of this research and point out its limitations and the possibilities for further research in this area. I shall then conclude by presenting the contributions it makes in both an academic perspective and a practical one.

## **1. 5. Structure of the Thesis**

The structure of this research is summarised in the research flow diagram shown in (Figure 1-3). This shows the outline of each chapter and the relations between them.



(Figure 1-3) Research Flow

## **CHAPTER 2. THEORETICAL BACKGROUND**

### **2. 1. Introduction**

Recently, most countries have conceived technological innovation as a very effective policy means to enhance national competitiveness and to improve the quality of life of citizens, so they have invested considerable national resources to carry out various types of R&D programmes. As a result, there has been a continuous need to evaluate the contribution of R&D activities to socio-economic development, the accountability of the expenditure on R&D fund and the quality of research performance at the national level (Ormala, 1989 : 333). For this reason, R&D evaluation has been considered one of the main research topics in both R&D management and policy evaluation in recent decade (OECD, 1983 ; Evered and Harnett, 1989 ; OECD, 1997). In other words, it is evident that the evaluation of S&T policy and large-scale R&D programmes in public sector has become a typical area of evaluation research.

In this chapter, I shall first enter into a brief and general discussion of policy evaluation and R&D evaluation that are slightly different in their historical development, as they are covered by different academic fields, although there has recently been active communication between them through the efforts of international organisations, including the OECD. However, the principal focus of my discussion will be on certain aspects of policy evaluation and R&D evaluation that show similarities and differences, rather than on overall theories relating to them. This chapter comprises three main sections. Section 2.2 relates to the definition of policy evaluation and its types, and then Section 2.3 deals with the definition of R&D evaluation, its types and major components. Finally, in Section 2.4, I shall address the relations between the two evaluation areas by suggesting some attributes and features of R&D evaluation that can be observed from the perspective of policy evaluation theory in general.

### **2. 2. Policy Evaluation**

Even though there have been many theoretical discussions concerning policy evaluation,



the emphasis of further discussion herein will be only on the definition of policy evaluation and its types. Through this discussion, I shall attempt to understand the definition and types of R&D evaluation from the perspective of policy evaluation theory, which are slightly different from those normally discussed in the area of R&D management.

### ***2.2.1. Definition of Policy Evaluation***

As in other areas of the social sciences, there is no single or simple definition of policy evaluation,<sup>1</sup> because many researchers have proposed their own definitions from different perspectives. In particular, there are many controversial debates about whether an evaluation can include the assessment of both ongoing and not-yet-implemented policy and about which types of results that the policy has realised can belong to evaluation (Vedung, 1997 : 7-8).

Regarding the first debate, there are some researchers who argue that prospective assessment, including *ex ante* assessment, forethought evaluation, needs assessment and analysis for goal setting, does belong to evaluation in a broader perspective (Anderson and Ball, 1978 : 3, 11 ; Rossi and Freeman, 1993 : 18). In this view, evaluation becomes an umbrella, covering all kind of analyses of public policy. The major argument against including *ex-ante* assessment in evaluation is drawn from the origin and history of evaluation research. The demands of the early evaluation movement for empirical data on policy and programme results emerged in opposition to the prevailing emphasis on the analysis of planned policy (Vedung, 1997 : 7-8). Therefore, it is rational that policy evaluation should be confined to after-the-fact assessments. Such evaluation concerns adopted policies in the sense of ongoing or terminated policies, programmes and programme ingredients. As a result, any prospective assessment can be classified as a “policy analysis” rather than a “policy evaluation”.

---

<sup>1</sup> Regarding the terms ‘policy evaluation’ and ‘programme evaluation’, there are some researchers who argue that the two terms have the same meaning. They argue that the difference between policy and programme is normally made by a relative standard rather than an absolute one and that, in fact, the conduct of policy evaluation comprises the execution of the evaluation of detailed programmes. Although there is good reason for this discussion in some cases, it is evident that the ultimate objectives of a policy and a programme are different in their level and scope. For this reason, in this thesis, the three terms, policy evaluation, programme evaluation and project evaluation are used with different meanings.

In relation to the second debate, it is sometimes argued that evaluation should be restricted to the results of ongoing activities, leaving out assessment of output or impact of finished policies and programmes from the narrow perspective (Nachmias, 1979 : 3f ; Rutman, 1980 : 17). There is no doubt that ongoing policies belong to the subject matter of evaluation. However, it is not appropriate that the core subject of policy evaluation is restricted to the process or results of ongoing policies. This is because careful assessment of terminated policies may very well assist in improving ongoing policies. Therefore, terminated policies and programmes may also need to be scrutinised for accountability. Consequently, it is rational that policy evaluation should include all retrospective or *ex-post* research of the administrative processes, immediate outputs, intermediate outcomes and long-term impacts of ongoing policies and programmes as well as terminated ones. (Lasswell, 1951 : 3-5 ; Vedung, 1997 : 8-9).

From the above discussions, we may infer that an expanded definition of policy evaluation has advantages over existing definitions. So, I shall propose the following broader definition of the policy evaluation suitable for the research purposes of my thesis : “policy evaluation is to retrospectively review and examine the implementation process of a certain policy, the types of immediate outputs that the policy has realised, and the mid and long-term influences of the policy outcomes and impacts on society as a whole as well as the overall system and process on which the policy has been based, so that it aims to contribute to the development of policy formulation and the improvement of policy implementation in future”. In accordance with the expanded definition, the areas of policy evaluation are shown in (Figure 2-1).

Object \ Timing	Before Implementation	During Implementation	After Implementation
Implementation Process	-	■	□
Policy Outputs	-	-	■
Policy Outcomes & Impacts	-	□	■
Overall Policy System & Process	-	■	■

■ : Primary Concerns of Policy Evaluation

□ : Secondary Concerns of Policy Evaluation

(Figure 2-1) Main Areas of Policy Evaluation

### 2.2.2. *Types of Policy Evaluation*

Many types of policy evaluation have been suggested depending on the main interest of researchers. In particular, the researchers who led evaluation research in the early stages proposed many types of policy evaluation by adopting different standards of classification. Glaser and Backer (1972 : 56-60) grouped these into three types : self-evaluation, internal evaluation and external evaluation by using the standard of those who are the evaluators. Scriven (1991 : 277-279) proposed *ex-ante* evaluation, process evaluation and *ex-post* evaluation as the main types of policy evaluation based on the time frame of evaluation. There are researchers who discuss six type of policy evaluation in accordance with the evaluation purposes as follows : endeavour evaluation, sufficiency evaluation, effect evaluation, efficiency evaluation, process evaluation and impact evaluation (Suchman, 1967 : 61-71 ; Nachmias, 1979 : 5-6). The Evaluation Research Society of the USA (1982 : 7-19) synthesised existing research concerning this subject and then suggested the following integrated six types of policy evaluation : just-before-implementation analysis, evaluability assessment, formative evaluation, effect or impact evaluation, programme monitoring and meta-evaluation.

From the above discussions, it appears that existing research includes the prospective or *ex-ante* assessment of not-yet-adopted policies as one type of policy evaluation. However, it is evident that the classification of types of policy evaluation should have a close connection with the definition of policy evaluation discussed above. Accordingly, I shall propose the following three types of policy evaluation : “process evaluation”, “summative evaluation” and “meta-evaluation”. In what follows, further brief explanation of these three types of policy evaluation will be given.

First, a process evaluation attempts to analyse through which path the inputs of a certain policy affected its results and what the reasons were for the results of the policy. In other words, it normally attempts to trace all kinds of policy consequences, including intended effects, null effects, perverse effects and side effects, whether advertent or inadvertent (McGraw et al., 1989 : 459-460 ; Scheirer, 1994 : 40 ; Vedung, 1997 : 210).

Since a process evaluation seeks to establish a whole pattern of causal interdependencies, it is important to identify the factors of the process of implementation that are supposed to

influence the policy outcomes. Regarding this matter, Vedung (1997 : 212-213) firstly discussed the basic issue of explanatory process evaluation in the following terms : (1) why is it that actual results – inside and outside the target area – of public policy may differ from the policy maker's initial expectations? ; and (2) why is it that desired outcomes are brought about without the aid of the policies, or even in spite of the policies? Then, he suggested that there were the following six detailed explanatory factors in process evaluation : (1) historical background of the public policy, (2) policy design, (3) implementation, (4) addressee responses, (5) other government policies and other government agencies and (6) issue networks and other environments.

Second, a summative evaluation seeks to judge whether a certain policy has succeeded or failed, through thoroughly investigating its results in some period after its implementation or termination (Hollister Jr. et al., 1979 : 140-158 ; Bryson and Cullen, 1984 : 267-289). A summative evaluation is considered to be an essential opportunity in the general policy process to provide policy makers with information to decide whether the pertinent policy will be expanded, modified or reduced and to know whether the related policy means need to be revised or changed. In addition, a summative evaluation sometimes functions to induce the conduct of a process evaluation. After the comprehensive examination of the overall results of the policy, there is also a need to investigate how the detailed path and process followed functioned to produce various results of the pertinent policy (Scriven, 1991 : 340).

In general, policy results can be divided into policy outputs and policy outcomes/impacts according to whether they will be realised immediately or in some period after completion of the policy and the extent to which they will directly or indirectly affect policy targets (Jenkins, 1998 : 31). In line with this classification of policy results, summative evaluation is also classified into “output evaluation” to evaluate immediate and direct outputs and “impact evaluation” to investigate relatively long-term and indirect impacts. Even though there has been a tendency to emphasise the direct and immediate policy outputs in summative evaluation, there is now a strong need to evaluate the mid and long term and sometimes, indirect, impacts of policy. This is because, since the influences of public policy on citizens and society as a whole continue to be deeper and stronger than in the past, we are now much interested in them.

Third, a meta-evaluation simply means the evaluation of another evaluation (Orata, 1940 ; Scriven, 1969). To be more precise, it is the evaluation of the findings, contents and processes of a certain evaluation during and after the evaluation activities, and of the evaluation system itself or structure in broader terms (Cook and Gruder, 1978 ; Stufflebeam, 1981 ; Larson and Berliner, 1983 ; Hoogerwerf, 1992 ; Rogers and Hough, 1995 ; Rogers, 1995). While over the last decades research has been conducted into meta-evaluation on a limited scale, it has recently become one of the promising subjects in the area of evaluation research (Georghiou, 1999 ; Horvat, 1999 ; Uusikylä and Virtanen, 2000 ; Curran, 2000).

Based on the above discussions, the types of policy evaluation for each main evaluation object are summarised in <Table 2-1>.

**<Table 2-1> Types of Policy Evaluation and Main Evaluation Object**

Type of Evaluation	Process Evaluation	Summative Evaluation		Meta-Evaluation
		Output Evaluation	Impact Evaluation	
<b>Main Evaluation Object</b>	<ul style="list-style-type: none"> <li>- Production path of policy result</li> <li>- Each stage of implementation process</li> </ul>	Primary result of implementation	<ul style="list-style-type: none"> <li>- Secondary results of implementation</li> <li>- Mid &amp; long-term impacts on society</li> </ul>	<ul style="list-style-type: none"> <li>- Findings, contents &amp; process of other evaluation</li> <li>- Evaluation system itself</li> </ul>

### 2. 3. R&D Evaluation

In general, R&D evaluation has been discussed as a part of R&D assessment. Therefore, it is argued that, although the underlying principles of R&D evaluation are broadly similar to those for other policy evaluations (HM Treasury, 1988 : 17), there are also not insignificant difference between policy evaluation in general and R&D evaluation. Further discussion will now be confined to a few elements of R&D evaluation, namely its definition, types and the major components necessary for its actual conduct, from which it is expected that the similarities and differences between the two areas of evaluation research will be identified.

### **2.3.1. Definition of R&D Evaluation**

As already indicated, the definition of R&D evaluation has tended to be discussed in close connection with that of R&D assessment. Regarding this subject, the Cabinet Office (1989 : 6) proposed that an R&D assessment is to provide the basis for informed decision making at all levels of management, in the initiation, selection, direction and termination of R&D activities. From this, it appears that the general definition of R&D assessment includes all activities before R&D start, during R&D implementation and after the completion of R&D. As already mentioned, in my thesis, the term evaluation relates only to retrospective judgement during or after implementation of a certain policy. Therefore, there is a need to slightly revise the existing definition of R&D evaluation to make it compatible with the research purposes of my thesis.

Meanwhile, until now, the purposes of R&D evaluation have been mainly limited to the internal utilisation of the results. However, recently, the focus of R&D evaluation has changed from the in-house efficiency of R&D activities and the internal utilisation of R&D results to the impacts of R&D in the technical, social and economic fields as a whole (Chung and Grupp, 1990 ; Geisler, 1994 ; Meyer-Krahmer, 1995 ; Georghiou, 1998 ; Georghiou and Roessner, 2000). On this basis, I shall define R&D evaluation as “activities to review and judge the R&D outputs or outcomes, and the diverse *ex post* impacts of R&D results over the entire societal system precisely and systematically during and after R&D activities”. Therefore, the appraisal of R&D activities, which is part of the general R&D assessment, is excluded from the subject of the R&D evaluation.

### **2.3.2. Types of R&D Evaluation**

Most studies have classified R&D evaluation into *ex ante* evaluation, interim evaluation and *ex post* evaluation on the basis of the time frame of evaluation (Gibbons and Georghiou, 1987 : 18-19 ; Cabinet Office, 1989 : 7 ; Nyden, 1992 : 166 ; Kostoff, 1994 : 189-190 ; Meyer-Krahmer, 1995 : 605 ; De Bant, 1995 : 371-372). Just as the results of policy implementation are divided into short-term policy outputs and long-term policy impacts (Jenkins, 1998 : 31), the results of R&D can be divided into short-term R&D outputs and the long-term impacts of R&D on the economic and social system. Hence, *ex*

*post* R&D evaluation can be reclassified into output evaluation and impact evaluation. From this perspective, I shall propose the classification of R&D evaluation into four types : “*ex ante* evaluation”, “interim evaluation”, “output evaluation” and “impact evaluation”. The individual types of R&D evaluation and their principal purposes are shown in <Table 2-2>. However, for the reason already discussed in relation to the scope of evaluation, *ex-ante* evaluation will be in principle excluded from further discussion, even though it is a type of R&D evaluation and, in particular, of R&D assessment.

<Table 2-2> Types of R&D Evaluation and its Principal Purposes

Type		Timing	Principal Purposes
<b><i>Ex-ante</i> Evaluation (Appraisal)</b>		Before R&D start	- selection of R&D programme/project - establishment of R&D objectives - budget allocation etc.
<b>Interim Evaluation (Monitoring/Review)</b>		During R&D implementation	- judgement of objectives in each stage - diagnosis of problems - revision or re-establishment of research objectives etc.
<b><i>Ex-post</i> Evaluation</b>	<b>Output Evaluation</b>	Just after R&D completion	- attainment of research objectives - examination of research outputs etc.
	<b>Impact Evaluation</b>	Some period after completion	- utilisation of research results - contribution to academy, industry, economy and social system etc.

### 2.3.3. Major Components of R&D Evaluation

When we try to judge R&D evaluation systems and their exercise, rational and appropriate components of them need to be selected. On this subject, Gibbons & Georghiou (1987 : 17-23) and Georghiou (1989 : 16) propose the following as essential components of R&D evaluation : evaluation purpose, evaluation object, evaluators, evaluation criteria, evaluation resources and evaluation audience. On the basis of these discussions, Lee et al. (1994a) and Lee & Suh (1996) also suggest slightly improved components of R&D evaluation.

From the synthesis of existing research, it is apparent that the current components of R&D evaluation need to be slightly re-classified in order to be compatible with the research purposes of my thesis. In particular, for an R&D evaluation to function as a

policy evaluation, the evaluation paradigm and the evaluation utilisation must be added to existing components of R&D evaluation. Just as a policy paradigm can prescribe the detailed directions, contents and characteristics of a specific policy, so in policy evaluation, what we can call an evaluation paradigm will prescribe the basic directions, contents and characteristics of specific evaluation activities (Park, 1996 : 40 ; Kim et al., 1995 : 164-165). The major elements of the evaluation paradigm will be the evaluation purpose, the evaluation object and the evaluation types (Yi and Kang, 2000b : 160). In addition, as one of the major goals of policy evaluation is to improve the policy processes and activities, the evaluation utilisation should be the most important of the evaluation components. The major elements of evaluation utilisation will be the evaluation report and the type of utilisation of evaluation findings (Kim, 1993 : 181-183, 194-195 ; Horvat, 1999 : 538).

Consequently, in order to correspond to the components of the meta-evaluation model that will be discussed in Chapter 3, the main components of R&D evaluation comprise (1) evaluation paradigm, (2) evaluation resources, (3) evaluation performance and (4) evaluation utilisation. These four main components can also be divided into many sub-components of R&D evaluation.

Of these sub components, some have great similarity to general components of policy evaluation, whereas others have their own particular attributes and features that have their origin in the different academic field of R&D management and are therefore very different from those of policy evaluation. In what follows, in order to incorporate the existing theoretical discussions about policy evaluation and R&D evaluation, further exploration will be devoted only to some individual components that are expected to be specific to the area of R&D evaluation. In principle, these include the following seven components : the evaluation purpose, the evaluation object, the evaluation type, evaluators, evaluation timing, evaluation criteria and evaluation methods. Despite this classification, some components that have already been dealt with in <Table 2-2>, namely the purpose of R&D evaluation, its types and timing, will also be excluded from the discussion. As a result, further discussion will focus only on the following four components of R&D evaluation : (1) evaluation object, (2) evaluators, (3) evaluation criteria and (4) evaluation methods.



(1) Object of R&D Evaluation : As in other policy evaluation, the most important aspect of R&D evaluation lies in determining the object to be evaluated. On this subject, most researchers have proposed that R&D evaluation should include the following objects : individual researchers, R&D projects, R&D programmes, research institutions and R&D policy (Gibbons and Georghiou, 1987 : 18 ; Georghiou, 1989 : 18 ; De Bant, 1995 : 371).

Regarding the selection of the evaluation object, there is a problem because the level in the research system at which it is desired to carry out the evaluation cannot always be easily isolated from the layers both above and below it. In particular, there is a problem in isolating R&D projects from R&D programmes as well as in isolating R&D programmes from R&D policy. Despite this problem, I shall propose the four following independent objects of R&D evaluation : R&D projects, R&D programmes, research institutions and R&D policy. Thus, individual researchers will be excluded from the research scope of my thesis, even though they are usually the object of R&D evaluation and, in particular R&D assessment. This is because the basic concept and methods of evaluating the performance of a particular researcher are different from those of evaluating the other four evaluation objects mentioned above (Cabinet Office, 1989).

(2) Evaluator of R&D Evaluation : The choice of who is to carry out the evaluation is very significant, because their objectivity and expertise will have a great deal of influence on the overall evaluation process, in particular, the final evaluation findings and their utilisation. Therefore, it is essential that the evaluators should be selected to achieve the initial evaluation purposes set up at the stage of evaluation planning (Horvat, 1999 : 535-536).

In general, the evaluators are classified as (1) internal evaluators and external evaluators and (2) peer reviewers and professional evaluators (Gibbons and Georghiou, 1987 : 11 ; Georghiou, 1989 : 20). While the internal evaluators belong to the organisation that has principal responsibility for the evaluation objects, the external evaluators normally come from outside, for example from academia, industry, other ministries or institutions and the media, and are expected to have no interest in actual evaluation objects. Meanwhile, peer reviewers are typically chosen by existing committees or their secretariats, whereas professional evaluators may be sited within an organisation or engaged under contract. The peer review process addresses scientific merit and occasionally offers views on

operational implementation. Peer reviews do not often address issues of the impact or achievement of strategic goals, which tend to be the preserve of professional evaluators. As a result, there may be a trade-off between them in relation to the expertise and the objectivity of evaluation exercises (Georghiou, 1989 : 20).

(3) Criteria of R&D Evaluation : The evaluation criteria relate to the standard of judgement of R&D activities. Therefore, it is evident that the actual criteria vary depending on the purposes of R&D evaluation and the weight given to them is also dependent upon the actual evaluation object.

There has been some important research, including *The Bromley Report* (1972) and *The National Science Foundation (USA) Proposal Evaluation Criteria* (1987), in relation to the evaluation criteria in the area of R&D management. In particular, Weinberg (1989 : 3-4) proposes two categories of evaluation criteria : “internal” criteria and “external” ones. Internal evaluation criteria arise from within the science itself and are basically criteria of efficiency. Thus one internal criterion is the competence of the performers, whereas another is whether the science was ripe for exploitation. On the other hand, where the required resources are large, internal criteria are insufficient. Therefore, external criteria must be considered. External criteria include the following criteria of utility : technological merit, social merit and scientific merit. That is, they measure the degree to which the given research, if successful, is, in the broadest sense, useful outside the field itself.

Even though described in many different ways in existing research, the evaluation criteria for R&D activities appear to be usually grouped into four broad categories : (1) those addressing the quality and originality of the scientific and technological contents, (2) those addressing the strategic goals of a programme or a related policy, (3) those concerned with the operational implementation and (4) those relating to effects and impacts of R&D activities on the economic and social systems (Georghiou, 1989 : 19).

(4) Method of R&D Evaluation : There is considerable research concerning evaluation methods, including the *ex-ante* evaluation of R&D (Gibbons and Georghiou, 1987 : 26-42 ; Chung and Grupp, 1990 : 164-165 ; Miller, 1992 : 3-7 ; Kostoff, 1994 : 190). Since *ex-ante* evaluation is excluded from the research scope of my thesis, further

discussion will focus on methods of *ex-post* evaluation of R&D activities. Therefore, it is essential that the evaluators should adopt the most suitable and applicable evaluation methods to achieve the principal evaluation purposes.

Geisler (1994) synthesised previous research and approaches in relation to the evaluation methods and proposed five categories of models, methods and techniques to evaluate R&D activities. She also discussed strengths and problems of individual evaluation methods. Therefore, it is apparent that her five categories of evaluation method will become very useful references for both evaluators and researchers in the area of R&D evaluation. The first category of evaluation method proposed by Geisler relates to the research on the “economic benefits” and economic assessment of the R&D process and its performers. The second category concerns the research that emphasises issues of “individual and group productivity” in the evaluation of R&D activities, in particular at project level. The third category includes evaluation efforts utilising “specific indicators”, some of which are quantitative, in the form of a single indicator or a combination of more than one. Notable in this category is research employing such indicators as publications, citations and patents. A fourth category of prior research deals with approaches “linking inputs to R&D with measures of outputs” from R&D. Such statistical correlations purport to indicate the contribution of R&D to economic growth and social entities. The fifth and last category includes evaluation approaches utilising subjective assessment by peers, clients and competitors. These techniques have been applied to the assessment of individual scientists, groups and entire R&D organisations (Geisler, 1994 : 190-191).

## **2. 4. Relations between Policy Evaluation and R&D Evaluation**

From the above discussions, it is evident that there are many similarities between policy evaluation and R&D evaluation, even though the academic field they originate from are different and the meaning of terms usually used in each field is also slightly different. In this connection, among the four types of R&D evaluation, both interim evaluation and *ex-post* evaluation functions are typical of policy evaluation. While the former attempts to thoroughly investigate the implementation process of R&D activities, the latter attempts to evaluate immediate outputs and long-term impacts of R&D after its completion. From this, it is clear that the interim evaluation of R&D activities corresponds to the process

evaluation of policy evaluation, whereas their *ex-post* evaluation is equivalent to the summative evaluation of policy evaluation.

However, *ex-ante* evaluation of R&D is expected to function as policy analysis in the area of policy sciences, even though it normally uses the term evaluation. This is because *ex-ante* evaluation emphasises the selection of programmes and projects in order to effectively achieve various objectives of S&T or R&D activities and then the budget allocation to them. In general, policy analysis can be defined as an intellectual and analytical assessment carried out prior to policy formulation in order to determine more rational policy objectives and to select the most effective policy means for the realisation of the policy objectives (Parsons, 1997 : 55 ; Jenkins, 1998 : 34-37). From this perspective, in the *ex ante* evaluation of R&D, while the selection of R&D programmes and projects equates to the determination of policy objectives, the budget allocation to them and the setting of their priority are considered as the investigation of more effective policy means.

<Table 2-3> Areas of Policy Evaluation and its Main Types

Timing Object	Before Implementation	During Implementation	After Implementation
Policy Objectives & Alternatives	<i>Policy Analysis</i>		
Implementation Process & Path		Process	Evaluation
Policy Outputs			Summative Evaluation (Output Evaluation)
Policy Outcomes & Impacts			Summative Evaluation (Impact Evaluation)

<Table 2-4> Areas of R&D Evaluation and its Main Types

Timing Object	Before R&D Start	During R&D Implementation	After R&D Completion
Selection & Budget Allocation	<i>Ex-ante Evaluation (Appraisal)</i>		
Implementation Process		Interim Evaluation	
R&D Outputs			Ex-post Evaluation (Output Evaluation)
R&D Outcomes & Impacts			Ex-post Evaluation (Impact Evaluation)

<Table 2-3> and <Table 2-4> are useful references for a more comprehensive understanding of the relations between policy evaluation and R&D evaluation. In two tables, the shaded cells represent the activities that are expected to be performed from the perspective of policy evaluation in general.

Consequently, *ex-ante* evaluation of R&D will be excluded from the scope of my thesis. Therefore, it will also not be dealt with in the case studies of the UK and the Korean national R&D evaluation systems and their meta-evaluations as well as in the suggestions for the applicability of some aspects of the UK case to the Korean case.

# **CHAPTER 3. RESEARCH FRAMEWORK**

## **: META-EVALUATION MODEL**

### **3. 1. Introduction**

Modern administrative nations have carried out evaluations of government activities in various forms for a long time. However, it has been revealed that there are sometimes large gaps between the ideal state and the reality of evaluation. This is due to the fact that, even though evaluators have the illusion that policy makers and policy implementers will take full account of their evaluation findings in the next stage of both policy formulation and policy implementation, this is not the usual case in actual policy processes (Barkdoll, 1980 : 174 ; Schneider, 1986 : 356). For this reason, evaluation researchers tend to be doubtful about the role of policy evaluation in policy process and the utilisation of evaluation findings. As a result, they also have suggested the need for research on the “evaluation of evaluation”, that is, meta-evaluation.

The need for meta-evaluation research first arose in the first half of the twentieth century in the area of educational evaluation (Orata, 1940). After that, there were few significant research results regarding this matter until the 1960’s. However, from the late 1960’s, with the development of policy evaluation research, there has been much theoretical discussion of meta-evaluation, as well as case studies which have evaluated previous evaluation activities and their outcomes in the areas of social programmes, educational programmes, R&D programmes and so on.

Accordingly, the aims of this chapter is to design a meta-evaluation model as the research framework that can be used in actually carrying out meta-evaluation, with particular reference to the area of R&D evaluation activities. I first intend to examine and analyse the existing literature directly or indirectly relating to meta-evaluation and, on this basis, to set up my working definition of meta-evaluation. Next, I will attempt to extract the components required in carrying out actual meta-evaluation, and to design a meta-evaluation model for this thesis based on these components.

### 3. 2. Review of Existing Research

There are two kinds of literature relating to meta-evaluation. One deals with this topic directly and explicitly referring to the term 'meta-evaluation' in the title or contents as a whole. On the other hand, there is another body of literature that deals with this subject in a more indirect, implicit way. The directly related literature is again divided into three types : the first involves theoretical discussion, the second includes theoretical discussion combined with case studies and the third type includes just simple case studies. These classifications are summarised in <Table 3-1>.

**<Table 3-1> Literature Regarding Meta-evaluation**

<b>Directly Related Literature</b>			<b>Indirectly Related Literature</b>
<b>Theoretical Discussion</b>	<b>Theoretical Discussion and Case Study</b>	<b>Case Study</b>	
<ul style="list-style-type: none"> <li>- Orata (1940)</li> <li>- Scriven (1969)</li> <li>- Cook (1974)</li> <li>- Cook &amp; Gruder (1978)</li> <li>- Stufflebeam (1981)</li> <li>- Larson &amp; Berliner (1983)</li> <li>- Schwandt &amp; Halpern (1988)</li> <li>- Greene (1992)</li> <li>- The Joint Committee (1994)</li> <li>- AEA (1995)</li> <li>- Rogers &amp; Hough (1995)</li> <li>- Rogers (1995)</li> <li>- Vedung (1997)</li> <li>- Praestgaard (1999)</li> <li>- EC (2000)</li> <li>etc.</li> </ul>	<ul style="list-style-type: none"> <li>- Apthorpe &amp; Gasper (1982)</li> <li>- Smith (1990)</li> <li>- Hoogerwerf (1992)</li> <li>- Georghiou (1999)</li> <li>- Horvat (1999)</li> <li>- Uusikylä &amp; Virtanen (2000)</li> <li>- Curran (2000)</li> <li>- Hong (2000)</li> <li>- Yi and Kang (2000b)</li> <li>etc.</li> </ul>	<ul style="list-style-type: none"> <li>- White et al. (1984)</li> <li>- Bobe (1991)</li> <li>- Gallegos (1994)</li> <li>- Georghiou (1995)</li> <li>- Hager and Hasselhorn (1998)</li> <li>- Cooksy (1999)</li> <li>- Grasso (1999)</li> <li>- Sanders (1999)</li> <li>etc.</li> </ul>	<ul style="list-style-type: none"> <li>- Barkdoll (1980)</li> <li>- Palumbo &amp; Nachmias (1983)</li> <li>- Strube &amp; Hartmann (1983)</li> <li>- Chelimskey (1987)</li> <li>- Nagel (1987)</li> <li>- The Joint Committee (1988)</li> <li>- Waller (1992)</li> <li>- Mackay (1992)</li> <li>etc.</li> </ul>

Meanwhile, with the revitalisation of meta-evaluation research, there have been case studies which have carried out meta-evaluation in diverse fields, such as social

programmes, educational programmes, R&D programmes, regional development programmes and so on. These have adopted the method of either evaluating the evaluation report or evaluating both the evaluation report and the field survey, as summarised in <Table 3-2>. However, because some case studies described here have examined only some components of meta-evaluation in broad terms, they cannot be regarded as a systematic and comprehensive meta-evaluation. Furthermore, most case studies have considered only a few components which have been simply derived from the basic concept of meta-evaluation, that is, the evaluation of evaluation, without any effort to design a meta-evaluation model suitable to both the characteristics of the programme meta-evaluated and the nature of the primary evaluation conducted. Therefore, it is apparent that the contribution of such case studies to the theoretical development of meta-evaluation is of limited significance.

**<Table 3-2> Types of Case Studies in Meta-evaluation**

<b>Method Used</b> <b>Object</b>	<b>Evaluation of Evaluation Report</b>	<b>Evaluation of Evaluation Report and Field Survey</b>
Social Programme	- Sanders (1999) - Grasso (1999) - Smith (1990)	-
Educational Programme	- Gallegos (1994) - Cooksy (1999)	- Hager & Hasselhorn (1998) - Curran (2000)
R&D Programme & Research Institute	- White et al. (1984) - Bobe (1991)	- Georghiou (1995) - Georghiou (1999) - Horvat (1999) - Hong (2000) - Yi & Kang (2000b)
Regional Development Programme	-	- Apthorpe & Gasper (1982) - Uusikylä & Virtanen (2000)
Sports Policy	- Hoogerwerf (1992)	-

As already indicated, the major purpose of this chapter is to determine the literature that is related to designing a research framework (meta-evaluation model) which can be used to evaluate the evaluation systems of S&T and R&D sector in the UK and Korea, and to identify the extent to which it is useful for this purpose. Thus, I will first confine the



literature review to the research articles and books which deal directly with the theoretical discussion of meta-evaluation. Of these, I will review only major research materials which have focused on theoretical aspects, such as methods, standards, criteria, principles or types of meta-evaluation, that are directly or indirectly relevant to the components of meta-evaluation.

### ***3.2.1. Models of Meta-evaluation in Cook and Gruder (1978)***

Cook and Gruder (1978 : 6), perhaps presenting the first systematic approach to meta-evaluation, use the term “meta-evaluation” to refer only to the evaluation of empirical summative evaluation, that is, where the data are collected directly from programme participants within a systematic design framework. Thus, their definition of meta-evaluation is narrower than that of Stufflebeam (1981) described below, whose discussion of meta-evaluation includes the evaluation of formative research and non-empirical evaluations (e.g. advocacy evaluation). Cook and Gruder also indicate that the purpose of meta-evaluation is simply to help evaluators meet their goals by providing diagnostic feedback and helpful advice about what to do.

As well as defining meta-evaluation, Cook and Gruder (1978 : 15-16) insist that the quality of past empirical summative evaluation appears to be low and explain the major reasons for this low quality in terms of the following eight problems : the failure of evaluation to consult formal decision makers, a vested interest of evaluation sponsors in the results, the low expertise of evaluation organisation, the negative conclusion of evaluations about impact, an inadequate design of evaluation research project, the lack of a proper monitoring system to detect problems early, deadline pressure and the difficulty in maintaining planned experimental contracts. Then, in order to improve the technical quality and relevance of evaluation research, particularly with respect to these problems, they suggest the seven meta-evaluation model shown in <Table 3-3>. They point out that, although these meta-evaluation models should help alleviate some of the problems, they are not by any means universal palliatives.

Cook and Gruder’s meta-evaluation model is generated on the basis of three variances. The first is whether the meta-evaluation takes a place simultaneously with the primary evaluation or after it ; the second is whether the primary evaluation data are or are not

manipulated by the meta-evaluator ; the third relates to the number of independent data sets that can be used to evaluate a particular programme. According to Cook and Gruder, this three variance framework stresses some of the more important attributes of useful evaluation. For example, the distinction between simultaneous and consecutive meta-evaluation highlights timeliness, with simultaneous studies being preferable to consecutive ones. The distinction between meta-evaluation with and without data manipulation stresses the enhanced validity and credibility that can result from the sophisticated manipulation of raw data by independent meta-evaluators. Also, the distinction between one and many data sets stresses the gains in credibility, validity and the range of questions that can be answered if there is a convergence of results across multiple data sets that evaluate the same programme (Cook and Gruder, 1978 : 17-18).

**<Table 3-3> Cook and Gruder's Meta-evaluation Models**

<b>Evaluation Timing</b>	<b>Data Manipulation</b>	<b>Data Number</b>	<b>Type of Meta-evaluation</b>
Subsequent to primary evaluation	Data not manipulated	Single data set	1. Essay review of an evaluation report
		Multiple data set	2. Review of the literature about specific programme
	Data manipulated	Single data set	3. Empirical re-evaluation of an evaluation or programme
		Multiple data set	4. Empirical re-evaluation of multiple data sets about the same programme
Simultaneous with primary evaluation	Data not manipulated	Single or multiple data set	5. Consultant meta-evaluation
	Data manipulated	Single data set	6. Simultaneous secondary analysis of raw data
		Multiple data set	7. Multiple independent replications

Source : Cook and Gruder (1978), pp.17.

Meanwhile, Cook and Gruder (1978 : 18-19) point out that these meta-evaluation models are not necessarily independent. For instance, if multiple independent replications of a

programme produced disparate findings, an integrative literature review would be required to try to make sense of them. Likewise, it is easy to imagine a consultant meta-evaluator obtaining copies of raw data and reanalysing them while the primary evaluator is also analysing them. Cook and Gruder also recall that, as in the case of all typologies, their meta-evaluation models are only justified by convenience rather than necessity.

In sum, Cook and Gruder's discussion on this subject focuses on improving the quality of summative evaluation, particularly with respect to the technical aspect of evaluation research. So, it is apparent that their research will make a great contribution to the development of a method in *ex-post* meta-evaluation at the level of individual policies or programmes. However, I think that their discussion is less relevant to both the methodological development in interim meta-evaluation and the design of framework for the evaluating an evaluation system at the national level.

### ***3.2.2. Roles of Meta-evaluation in Stufflebeam (1981)***

Stufflebeam (1981 : 150-151), who was the chairman of the Joint Committee on the Standards for Educational Evaluation that developed the *Standards for Evaluations of Educational Programs, Projects and Materials (1981)*, first explains that, since meta-evaluation is an instance of evaluation, the definitions of these two terms must be consistent. According to his explanation, given that evaluation is the assessment of the worth and merit of a certain programme and policy, meta-evaluation can be defined as the assessment of the worth and merit of evaluation. However, in contrast to Cook and Gruder, he suggests that it is desirable to extend the meaning of meta-evaluation to give some idea of what is meant by the worth and merit of an evaluation. He first comprehensively examines existing research relating to meta-evaluation. Next, he classifies this into four types : theoretical development processes, meta-evaluation performance examples, development processes for evaluation techniques useful to meta-evaluation and critical research into existing evaluation. After that, he advocates the expansion of existing concepts in meta-evaluation. He further proposes the following operational definition of meta-evaluation : a process of delineating, obtaining and using descriptive and judgemental information about the utility, practicality, ethics and technical adequacy of an evaluation in order to guide the evaluation and publicly to report

its strengths and weaknesses (Stufflebeam, 1981 : 147-151).

On the basis of this definition of meta-evaluation, Stufflebeam (1981 : 151-152) insists that its two main roles are to guide evaluation activities and to publicly report their strengths and weaknesses. The former role is formative meta-evaluation and the latter is summative meta-evaluation. He also insists that, of the two, summative meta-evaluation is the fundamental role of meta-evaluation.

Summative meta-evaluation sums up the overall merit of an evaluation and is usually done following the conclusion of a primary evaluation. It holds evaluators accountable by publicly reporting on the extent that their evaluation reports meet standards of good evaluation practice. Summative meta-evaluations also help the audiences of primary evaluations determine how seriously they should take those evaluation's reported conclusions and recommendations. In contrast, formative meta-evaluation is a more proactive and constructive application of evaluation. More detailed, formative meta-evaluations help evaluators to make decisions about how to conceive, plan, conduct, interpret and report their studies. It monitors the process of evaluation research and provides continuous feedback intended to help assure the quality of the evaluation. Formative meta-evaluation also has its foundations in evaluation guidelines and pitfalls that are believed to be linked to the meeting of the standards that serve as the basis for summative meta-evaluation. Therefore, the main assumption underlying formative meta-evaluation is that the ultimate value of an evaluation is maximised when certain procedural guidelines are met and pitfalls avoided. In other words, summative meta-evaluation is enhanced by paying attention to an appropriate, ongoing formative one (Stufflebeam, 1981 : 151-152).

Consequently, Stufflebeam makes a quite large contribution to the development of meta-evaluation in that he proposes a broader and more comprehensive definition of it and points out that it has two roles : formative and summative meta-evaluation. Since he extends the scope of meta-evaluation from the summative approach so that it also includes a formative one, it is apparent that his research will be a very useful reference for my thesis. However, it has the small limitation that it deals mainly with evaluation research in the area of educational programmes or projects. This means that some modification of Stufflebeam's theory will be necessary, before it is incorporated into the

design of my research framework.

### ***3.2.3. Components of Evaluation in Larson and Berliner (1983)***

Larson and Berliner (1983 : 154 ; 162), perhaps the first researchers to discuss the components of meta-evaluation systematically, first point out that, just as evaluations are performed to provide decision makers with information about a programme, so evaluations of evaluations (EOE's or meta-evaluations) are performed to provide information about an evaluation to a possible different set of decision makers. They further indicate that evaluations of evaluations (meta-evaluations) have several potential purposes, such as to provide an independent assessment to decision-makers of the quality of information contained in an evaluation, to provide guidance in selecting an evaluator, to assimilate research knowledge from a number of separate but similar programmes and to provide a vehicle for examining the evaluation enterprise itself. Therefore, the decision to perform a meta-evaluation is also an allocation of resources which may or may not be justified at a particular point in time, in terms of the marginal cost and expected marginal information content of the meta-evaluation, given one's knowledge about the original evaluation and the programme being evaluated.

On the basis of the above discussion, they insist that, just as any evaluation is a process which has inputs and yields outcomes, so any comprehensive meta-evaluation (EOE) should examine all three evaluation components : evaluation inputs, the evaluation process and evaluation outcomes. They also present an initial listing of the detailed elements that evaluation inputs, the evaluation process and evaluation outcomes comprise, which is shown in <Table 3-4>. Evaluation inputs may be considered to be an inventory of resources and methodologies that have a bearing on the evaluation and the basic elements of the evaluation and programme setting. The utilisation of evaluation inputs is an evaluation process, the actual conduct of the evaluation as compared with what was planned in the evaluation design. Finally, evaluation outcomes are the decisions (resource allocations) influenced by the evaluation (Larson & Berliner, 1983 : 155 ; 157 ; 159). Of course, we have to remember that these detailed lists of elements in each component of evaluation are illustrative and sometimes overlapping, not comprehensive and exclusive ones.

**<Table 3-4> Larson and Berliner's Evaluation Components**

<b>Evaluation Components</b>	<b>Detailed Factors for Evaluation</b>
Evaluation Inputs	<ol style="list-style-type: none"> <li>1. Budget and other material resources available to the evaluators</li> <li>2. Duration</li> <li>3. Timing with respect to the programme being evaluated</li> <li>4. Attributes of evaluation personnel (e.g., training, experience, world view, etc.)</li> <li>5. Attributes of programme personnel (e.g., experience, commitment, education, etc.)</li> <li>6. Programme attributes (e.g., goals, substantive area of concern, client group, etc.)</li> <li>7. Evaluation methodology and design</li> <li>8. Audience or client group, and purposes of the evaluation</li> <li>9. Existing data and data limitations</li> <li>10. Underlying theoretical model(s)</li> <li>11. The expected policy consequences</li> </ol>
Evaluation Process	<ol style="list-style-type: none"> <li>1. Types, intensity and frequency of interactions between evaluators and programme staff members</li> <li>2. Response of programme staff and client groups to the presence of evaluators</li> <li>3. Extent to which acquired information is fed back to programme staff, perhaps modifying programme procedures</li> <li>4. Extent to which acquired information is used to modify the allocation of evaluation resources</li> <li>5. Adaptiveness of evaluation design (e.g., capacity to respond to change in the programme)</li> <li>6. Changes in personnel (e.g., evaluators, programme staff, client groups of both programme and evaluation)</li> <li>7. Methodology : the formal and informal processing of information leading to evaluative findings</li> <li>8. Communication of findings</li> </ol>
Evaluation Outputs	<ol style="list-style-type: none"> <li>1. Decision by funding agency to fund, refund, modify or cancel programme</li> <li>2. Decision by programme staff to modify any of the programme procedures</li> <li>3. Decision by members of the client group to alter participation patterns in the programme</li> <li>4. Decision by one or more members of the research community to study further the questions/issues raised in the evaluation</li> <li>5. Decision by one or more other funders and/or programme personnel to initiate, modify, or terminate similar programmes</li> </ol>

Source : Larson and Berliner (1983), pp.155 ; 157 ; 160.

In sum, the components of evaluation proposed by Larson and Berliner will be a very useful reference for designing the research framework of this thesis, in that they identify

the main stages of evaluation activity as inputs, the process and outputs, and establish detailed factors for evaluation in each stage. However, they are not concerned with the evaluation purposes and evaluation objects which are more important in policy evaluation than the above-mentioned factors. Also, they are not greatly concerned with the rational and due procedures which are necessary to guarantee the transparency and independence of evaluation activities.

Hence, even though Larson and Berliner's proposals represent a significant development in meta-evaluation research, they have some limitations. (1) Firstly, as the components of evaluation inputs and the evaluation process mostly discussed by them, are instrumental and methodological factors, discussions regarding the evaluation purposes and evaluation objects which prescribe the directions and characteristics of the overall evaluation activities of specific policies or programmes are avoided. If the evaluation purposes and evaluation objects are not correctly set up in policy evaluation in advance, the evaluation findings may not be compatible with the paradigms and objectives of the individual policy or programme being evaluated, even though a competent evaluator may carry out the evaluation by using both large amounts of resources and diverse evaluation methods. Therefore, it is apparent that the discussion of these two factors is extremely important. (2) The second limitation is regarding the components of evaluation outcomes. Since discussions are focused on the direct feedback of evaluation findings in the next stage of the policy process, they appear to confine the utilisation of evaluation outcomes to instrumental utilisation. Although the normal way in which evaluation is utilised is to terminate a policy, make a policy change or improve a policy implementation, it is not appropriate that evaluation utilisation should merely be confined to the direct feedback of evaluation findings, since it is often the case that policy evaluation aims to produce information that is useful and relevant in the policy process as a whole. (3) Thirdly, while Larson and Berliner refer to 23 detailed evaluation factors in meta-evaluation, these factors sometimes overlap and their classification is slightly inappropriate. For example, it would be more appropriate if "users of evaluation outcomes" and "expected policy improvement potential", which are included in evaluation inputs, and "examination of evaluation report" which is included in the evaluation process were classified as evaluation outcomes. In addition, while "evaluation methodology" appears in both evaluation inputs and the evaluation process, it seems more proper to classify them into a single factor in the evaluation process. (4) A fourth limitation of Larson and Berliner's article is that there

is no evaluation of procedures. In order to obtain rational and adequate evaluation findings, evaluation procedures themselves should be implemented in accordance with a rational and due perspective, in addition to the components of evaluation inputs and the evaluation outcomes being appropriate. Consequently, in order to carry out a comprehensive and systematic meta-evaluation, both substantial and procedural factors relating to the evaluation performance must be simultaneously examined and judged.

### ***3.2.4. The Program Evaluation Standards in The Joint Committee (1994)***

In 1994, The Joint Committee on Standards for Educational Evaluation which represented 16 organisations in the area of evaluation, including the American Educational Research Association, the American Evaluation Association and the Canadian Evaluation Society, published *The Program Evaluation Standards : How to Assess Evaluations of Educational Programme (2<sup>nd</sup> Edition)*. In this book, the Joint Committee (1994 : 2) defines a standard as a principle mutually agreed to by people engaged in a professional practice, that, if met, will enhance the quality and fairness of that professional practice, for example, evaluation. It also describes the four roles of these standards as follows : (1) *The Standards* provide a guide for evaluating educational and training programmes, projects and materials in the variety of settings. (2) They provide a framework for designing and assessing the previous evaluations. (3) They encourage the use of a variety of evaluation methods. (4) They also help evaluators identify and confront political reality (The Joint Committee, 1994 : 1-4).

The Joint Committee developed 30 principles designed to assist both formative and summative meta-evaluation. They are grouped under the four main headings of utility, feasibility, propriety and accuracy, as shown in <Table 3-5>.

The Joint Committee believes that these four attributes are necessary and sufficient for sound and fair evaluation. (1) Utility standards guide evaluations so that they will be informative, timely and influential. They require evaluations to acquaint themselves with their audience, define the audience clearly, ascertain the audience's information needs, plan evaluations to respond to these needs and report the relevant information clearly and in a timely fashion. Overall, these standards define whether an evaluation serves the practical information needs of a given audience. (2) Feasibility standards recognise that



evaluations are usually conducted in a natural, as opposed to a laboratory, setting, and also consume valuable resources. Therefore, these standards call for evaluations to be realistic, prudent, diplomatic and economical. (3) Propriety standards reflect the fact that evaluations affect many people in a variety of ways. Such standards are intended to facilitate protection of the rights of individuals affected by an evaluation. (4) Accuracy standards determine whether an evaluation has produced sound information. Thus, they are intended to ensure that an evaluation will reveal and convey accurate information about the programme's merit and/or worth (The Joint Committee, 1994 : 5-6).

**<Table 3-5> The Joint Committee's Evaluation Standards**

<b>Standards</b>	<b>Principles</b>
Utility Standards	U1 Stakeholder Identification U2 Evaluator Credibility U3 Information Scope and Selection U4 Values Identification U5 Report Clarity U6 Report Timeliness and Dissemination U7 Evaluation Impact
Feasibility Standards	F1 Practical Procedures F2 Political Viability F3 Cost Effectiveness
Propriety Standards	P1 Service Orientation P2 Formal Agreements P3 Rights of Human Subjects P4 Human Interactions P5 Complete and Fair Assessment P6 Disclosure of Findings P7 Conflict of Interest P8 Fiscal Responsibility
Accuracy Standards	A1 Programme Documentation A2 Context Analysis A3 Described Purposes and Procedures A4 Defensible Information Sources A5 Valid Information A6 Reliable Information A7 Systematic Information A8 Analysis of Quantitative Information A9 Analysis of Qualitative Information A10 Justified Conclusions A11 Impartial Reporting A12 Meta-evaluation

While an important contribution to the development of evaluation and meta-evaluation in the area of the educational programme, these standards do not in themselves constitute a satisfactory framework for meta-evaluation. For a start, their emphasis on education and North American origins mean that they need to be checked for applicability in other countries and other types of policy or programme. Most importantly, they are designed to evaluate individual evaluations rather than evaluation approaches or evaluation systems. While they provide criteria for evaluating evaluations, they offer little guidance on how to plan evaluations in order to meet these standards. In fact, even though The Joint Committee suggests five steps for applying these standards to an actual evaluation, the suggestion of using these steps is itself an approach to evaluation, which requires evaluation and evidence to be provided of its usefulness (Rogers, 1995 : 4).

### ***3.2.5. Principles for Evaluators in the AEA (1995)***

In 1995, the American Evaluation Association (AEA) published *Guiding Principles for Evaluators*, the draft of which had been presented to and approved by a membership vote at the January 1994 AEA board meeting. The AEA's five principles are intended to foster the primary aim of evaluation : that is, for evaluators to aspire to conduct and provide the best possible information that might bear on the value of whatever is being evaluated. These principles proposed by the AEA are not independent, but overlap in many ways. They are also sometimes in conflict, so that evaluators will have to choose among them. An important consideration is that these principles are not intended to be or to replace standards supported by evaluators or by the other disciplines in which evaluators participate (AEA, 1995 : 20-21).

*Guiding Principles for Evaluators* of the AEA comprise five main principles each of which includes several sub principles, as follows (AEA, 1995 : 22-26) :

- (1) Systematic Inquiry : Evaluators should conduct systematic, data-based inquiries about whatever is being evaluated. They should :
  - adhere to the highest appropriate technical standards.
  - explore with the client shortcomings and strengths of both evaluation question and approaches to answering them.
  - communicate methods and approaches accurately and in sufficient detail, to make limitations clear, discuss values, assumptions, theories, methods, results

and analyses that significantly affect the interpretation of the findings.

(2) Competence : Evaluators should provide competent performance to stakeholders.

- Evaluation team should possess the appropriate education, abilities, skills and experience.
- Evaluators should practise within the limits of the professional training and competence.
- Evaluator should maintain and improve competence.

(3) Integrity/Honesty : Evaluators should ensure the honesty and integrity of the entire evaluation process. They should :

- negotiate honestly with clients and relevant stakeholders on costs, tasks, limitations of methodology, scope of results and the uses of data.
- record all changes made in the originally negotiated project plans, reason for them, inform client and other important stakeholders in timely manner.
- determine their own, their client's and other stakeholder's interest in the conduct and outcome of the evaluation, including financial, political and career interests.
- disclose any roles or relationships concerning what is being evaluated that might pose a significant conflict of interest.
- not misrepresent procedures data or findings.
- communicate likely misleading evaluative information to the client, and decline to proceed if client does not resolve concerns.
- disclose all source of financial support and of the request for an evaluation.

(4) Respect for people : Evaluators should respect the security, dignity and self-worth of respondents, programme participants, clients and other stakeholders with whom they interest. They should :

- abide by current professional ethics and standards regarding (a) the risks, harms and burdens on those participating in the evaluation, (b) informed consent and (c) the scope and limits of confidentiality.
- maximise benefits and minimise harm, without compromising the integrity of the evaluation.
- respect stakeholder's dignity and self-worth.
- foster social equity of the evaluation, and inform participants if and how they can receive services to which they are otherwise entitled, where feasible.
- identify and respect differences among evaluation participants in culture, religion, age, sexual orientation and ethnicity.

(5) Respect for general public welfare : Evaluators should articulate and take into account the diversity of interests and values that may be related to the general and public welfare. They should :

- include the important perspectives and interests of the full range of stakeholders.
- consider broad assumptions, implications and potential side effects, not only immediate operations and outcomes.
- allow all relevant stakeholders to have access to evaluative information, and actively disseminate that information.
- maintain a balance between client needs and other needs.
- go beyond an analysis of stakeholder interests when considering the welfare of society as a whole.

The AEA's principles also have certain limitations just like other standards or criteria for good evaluation. The AEA itself points out that these principles were developed in the context of Western cultures, particularly that of the United States, and so may reflect the experiences of that context. Thus, the relevance of these principles may vary in other cultures, and even across subcultures within the United States (AEA, 1995 : 21-22).

Meanwhile, there has been research to integrate and consolidate the two sets criteria while applying them in the actual evaluation activity (Sanders, 1995 ; Grasso, 1999), because the Joint Committee's *The Programme Evaluation Standards* and the AEA's *The Principles for Evaluators* take very different approaches to organising their criteria, potentially complicating the task. Grasso (1999 : 356), firstly, examines the two sets of criteria to determine whether and how they might be consolidated. After that, he concludes that each of *the Standards* could be subsumed into one of *the Principles*, in most cases relatively straightforwardly. Indeed, to some extent, the standards proved useful in providing more specificity in relation to a number of the principles, allowing a fuller discussion (Grasso, 1999 : 367-370).

### **3.2.6. Framework for Evaluating Evaluation in Rogers (1995)**

In her paper, Rogers (1995 : 1) firstly asks how we should evaluate the programme evaluation. Then, she insists that, while technical competence and utilisation of evaluation have often been used as the criteria for evaluating evaluation, they are not

sufficient for this purpose. After reviewing the existing research in the area of policy and programme evaluation, she also concludes that previous evaluation theory and practice has emphasised one or more of the following five criteria for evaluating evaluation : (1) producing valid information, (2) producing useful information, (3) producing an unbiased judgement of merit or worth of primary evaluation, (4) the participation and illumination of stakeholder and (5) empowerment of programme clients.

Rogers (1995 : 4-6) goes on to suggest a framework for evaluating evaluations (meta-evaluation) in terms of their intended and actual impact on the implementation of existing programmes. This framework consists of three major factors. These are (1) the five evaluation criteria for the intermediate outcomes of programme evaluation referred to above, (2) the evaluation's impact on the implementation of existing programmes or on their replacement with alternatives and (3) the evaluation's contribution to the development of programmes which can meet a certain policy need.

It is apparent that this framework is systematic and comprehensive, but that its primary objectives are limited to the evaluation of programmes that are being implemented at the moment. As a result, it seems that some small adaptations need to be made in order for Rogers' meta-evaluation framework to be adopted directly in a research framework for evaluating evaluation systems as a whole at the national level.

### **3.3. Definition of Meta-evaluation**

#### ***3.3.1. Review of the Existing Definition of Meta-evaluation***

While there is a lot of research on meta-evaluation, a simple definition is "evaluation of evaluation" (Cook and Gruder, 1978 ; Stufflebeam, 1981 ; Larson and Berliner, 1983) or "secondary evaluation" (Cook, 1974). However, just as the existing definition of meta-evaluation is not only fairly simple but also partial, so it is sometimes difficult to enter into an in-depth discussion regarding it. Therefore, in order to overcome these limitations, I intend to propose a working definition of meta-evaluation for my thesis, which is more comprehensive and systematic.

Even though policy evaluations are carried out for a variety of purposes, there are some cases where evaluation findings are not sufficiently utilised in reality. If a policy evaluation cannot achieve the objectives initially intended, this means that there is a problem in the evaluation itself. This makes it necessary to re-evaluate the rationality and appropriateness of evaluation activities themselves. Orata (1940) calls this type of evaluation an "evaluation of evaluation". Dror (1971) also stresses an "evaluation of evaluation" to promote the evaluation utilisations. Scriven (1969) creates the term "meta-evaluation" to refer to such secondary evaluation activities. As already indicated, Stufflebeam (1981 : 147-151) proposes an operational definition of meta-evaluation which has a more comprehensive concept than previous researches. According to him, meta-evaluation is a "process of delineating, obtaining and using descriptive and judgemental information about the utility, practicality, ethics and technical adequacy of an evaluation in order to guide the evaluation and publicly to report its strengths and weaknesses".

Though meta-evaluation is defined as above, individual scholars have made the following suggestions about its scope and objects : (1) Chelimsky (1985 : 11) proposes that, in the narrowest sense, meta-evaluation is understood as an evaluation synthesis of evaluation findings. Evaluation synthesis means the re-analysis of facts discovered in one or more evaluations, in order to examine what has been learned through a certain policy or programme. (2) Cook and Gruder (1978 : 60) suggest that meta-evaluation, in a narrow sense, means only the evaluation carried out on evaluation itself, and is limited to the empirical evaluation performed by directly collecting data from programme participants within a systematic design framework. (3) In contrast to the two types of definition, Stufflebeam (1981 : 150-152) proposes the extension of the definition to include formative meta-evaluation as well as summative meta-evaluation. Larson and Berliner (1983 : 148) also understand meta-evaluation as the evaluation of overall evaluation procedures, including evaluation inputs, the evaluation process and evaluation outcomes. (4) Hoogerwerf (1992), Rogers and Hough (1995) and Rogers (1995) propose the definition and scope of meta-evaluation in the broadest sense. That is, they include evaluation systems or evaluation structures as well as evaluation findings and the evaluation process in the scope of the meta-evaluation.

### 3.3.2. *New Definition of Meta-evaluation*

To overcome the partial and fragmental approach that characterise the definition of meta-evaluation in some previous research, and to offer higher generality and utilisation, it is necessary to define it in accordance with its original function. When we think that meta-evaluation was initially suggested to improve the quality of evaluation and to promote the utilisation of its findings, it is desirable to prescribe the definition and scope of meta-evaluation in relation to this concept. In addition, so that even policy evaluation itself, which is one of the management processes in the public sector, can be evaluated just like other policy processes, this chapter attempts to provide a new definition of meta-evaluation in a wider sense. The following is in fact proposed : “a special type of evaluation carried out to promote the utilisation of evaluation findings and to enhance the evaluation quality, by evaluating the evaluation paradigm, the evaluation resources, the evaluation performance and evaluation utilisation during and after the evaluation activities, and of the evaluation systems or structures in broader terms”.

The following list of the characteristics will throw further light on this slightly different definition of meta-evaluation.

- (1) The object of meta-evaluation is both formative and the summative evaluation which represent one factor of policy evaluation itself as a whole. In this respect, while meta-evaluation is one type of policy evaluation, it is different from a common policy evaluation, which evaluates specific policies and programmes.
- (2) Like a common policy evaluation, meta-evaluation can also be carried out during or after primary evaluation. For this reason, it has certain characteristics similar to a policy evaluation, that is, an *ex post* and retrospective activity which fundamentally discriminates the policy evaluation from other policy activities, in particular, policy analysis.
- (3) Meta-evaluation covers all stages of evaluation processes and of evaluation systems, not just specific aspects of evaluation activities. In other words, the scope of meta-evaluation normally covers not only the evaluation paradigm that can determine the direction of policy evaluation as a whole, but also the utilisation of evaluation findings that will often happen after its completion.
- (4) Meta-evaluation attempts to evaluate not only the content and results of evaluation activities, but also the evaluation procedures. This means that it is also interested in

the appropriateness of the content of policy evaluation as well as in its procedural rationale.

- (5) Meta-evaluation is performed to promote the utilisation of evaluation findings, by increasing its quality. Hence, one of its major purposes is not just to evaluate the evaluation activities, but rather to achieve the practical purpose of improving the utilisation and quality of evaluation themselves.

### **3. 4. Components of Meta-evaluation**

#### ***3.4.1. Limitations of Existing Research***

Just as the basic components of evaluation are required to carry out a policy evaluation, so the components of meta-evaluation must also be set up in detail in order to execute a meta-evaluation. Scriven presented the following eight kinds of components of policy evaluation : an evaluation purpose, an evaluation organisation or evaluator, a policy definition, an evaluation process, policy outcomes, policy impacts, evaluation criteria and evaluation utilisation (Shadish Jr. et al., 1991 ; Vedung, 1997 : 93). Even though evaluation methods are not referred to, it seems that these eight components do contain the factors for carrying out a policy evaluation. Therefore, such components of policy evaluation will be a very useful reference for deriving the components of meta-evaluation.

Meanwhile, as already indicated in Section 3.2, there is some existing research which deals directly or indirectly with this subject, covering such factors as components, standards, principles and the framework of meta-evaluation. However, this research has had certain limitations which make it difficult to adapt directly to the design of a meta-evaluation model. The limitations of the research in this existing literature can be summarised as follows :

- (1) Cook and Gruder's discussion is less relevant to both the methodological development of interim meta-evaluation and the design of a framework for evaluating the evaluation system at the national level, because it refers only to the evaluation of empirical summative evaluation.
- (2) Both Stufflebeam's theory and the Joint Committee's Standards need some



modification to be incorporated into the design of my research framework, because they deal mainly with evaluation research in the area of educational programmes or projects.

- (3) In Larson and Berliner's article, there are some shortcomings such as concentration on the instrumental and methodological factors, the omission of procedural factors in the evaluation performance and inappropriate classification of the components.
- (4) Since the AEA's principles were developed in the context of the USA, the relevance of these principles may vary in other cultures, such as those of European and Asian countries.

### ***3.4.2. Discussion of New Components in Meta-evaluation***

To overcome the limitations described above, this chapter will discuss new components of a more systematic and comprehensive meta-evaluation, by proposing the evaluation paradigm as new component, by expanding the concept of evaluation utilisation and by readjusting the classification of the existing proposed evaluation factors. In particular, the reason for discussing the evaluation paradigm as a new component of meta-evaluation is that, in policy evaluation, only when directive and paradigmatic factors have been set up correctly in advance, can instrumental factors such as evaluation resources and evaluation performance be rationally implemented. For this reason, in this chapter, I intend to classify the components of meta-evaluation into four types : (1) the evaluation paradigm, (2) evaluation resources, (3) evaluation performance and (4) evaluation utilisation. In what follows, the major sub components of the four main components will be extracted and discussed.

#### ***3.4.2.1. Evaluation Paradigm***

The concrete directions, contents and characteristics of specific policies are in reality determined by a policy paradigm. We can consider ideology, philosophy, principles, fundamental rules, theory and values as the main components of such paradigms (Park, 1996 : 40 ; Kim et al., 1995 : 164-165). Similarly, in policy evaluation, we can also suppose that there are roles similar to those of policy paradigms, which fundamentally prescribe the directions, content and characteristics of overall evaluation activities. In line

with this, the factors that will determine the directions and contents of policy evaluation as a whole at national level are the rationales for the evaluation purpose, the suitability of the evaluation objects and the appropriateness of evaluation types. This is because these components determine the input of evaluation resources, the contents of evaluation performance and the utilisation types of evaluation findings which are carried out after setting the evaluation paradigm (Yi and Kang, 2000a : 4-5). Therefore, since it can be assumed that evaluation purposes, evaluation objects and evaluation types play roles of evaluation paradigm in policy evaluation, their rationality and appropriateness must be firstly examined and judged in meta-evaluation.

### **(1) Evaluation Purposes**

In order to carry out meta-evaluation, the appropriateness of the evaluation purposes and their rational application should be evaluated. In other words, we must evaluate whether evaluation purposes are rationally established or not, and whether they can be compatible with the policy paradigms and policy objectives as a whole.

In general, while policy evaluations are carried out for rational and justifiable purposes, they are sometimes executed for such undesirable purposes as political advocacy. Most researchers suggest accountability, the improvement of policy/programme and the enhancement of basic knowledge as the rational purposes of policy evaluation (Goldenberg, 1983 ; Arvidsson, 1986 : 627 ; Lee and Sampson, 1990 : 157 ; Hudson, Mayne and Thomlison, 1992 : 5 ; Vedung, 1997 : 101). Less rational purposes of policy evaluation include the posture of policy effects, the postponement of decision making, the ducking of responsibility, public relations and the eyewash of policy effects (Weiss, 1972 : 11-12 ; Suchman, 1972 : 81). As already discussed, evaluation purposes will affect the whole process of evaluation, which ranges from the input of evaluation resources to the utilisation of evaluation findings. Thus, the meta-evaluator must accurately examine and judge for what purposes evaluators and managers have carried out a policy evaluation, and how they have attempted to apply these purposes to the actual evaluation.

### **(2) Evaluation Objects : Scope and Level**

To decide the scope and level of evaluation objects is the initial step in the actual evaluation as well as a very important activity that will determine its success or failure. If the scope of evaluation objects is excessively wide, it will be difficult to complete the

evaluation within the arranged time. However, if its scope is too narrow, the evaluation findings may cover only partial and non-essential issues. Meanwhile, if the evaluation objects are set up at an extremely low level, the negative perspectives of the evaluation such as one of management tools may appear. On the other hand, if the evaluation objects are determined at a relatively higher level, the evaluation findings may lack concreteness since they will not represent the real world. Meta-evaluators, therefore, should evaluate whether evaluation objects have been selected at a scope and level that can be controlled by the evaluators and can achieve the evaluation purposes.

### **(3) Evaluation Types**

The meta-evaluator needs to investigate whether evaluation types are rationally adopted in order to properly realise the evaluation purposes set up initially. In general, the policy evaluation tends to be carried out within the restricted resources so that all types of the evaluation cannot be conducted simultaneously. For this reason, the evaluators are always faced with the difficulty of adopting the evaluation types capable of most effectively achieving evaluation purposes selected in advance. For example, when evaluation purpose will be placed on the accountability of the related institutions or staff concerned, the adoption of an output evaluation will be more effective. On the other hand, if the improvement of the policy or programme would be emphasised, an interim evaluation or a process evaluation will well function. Meanwhile, an impact evaluation can effectively contribute to the realisation of a long-term evaluation purposes, including the knowledge enhancement.

#### ***3.4.2.2. Evaluation Resources***

Evaluation resources comprise the variety of tangible and intangible elements which are used during the actual evaluation. Since the quantity and quality of evaluation resources can seriously affect the quality of evaluation performance and evaluation findings, it is a quite important to examine them in meta-evaluation. According to their characteristics, resources for the policy evaluation can be classified into the categories of evaluation manpower, the evaluation unit and the evaluation information.

##### **(1) Evaluation Manpower : Quantity and Quality**

It is quite apparent that the quality of the evaluation findings is determined by the quantity

and quality of the evaluation manpower. Therefore, in order to ensure an effective policy evaluation, first of all, there must be sufficient evaluation manpower. However, it is not enough to simply count the number of evaluators when we evaluate the quantitative aspect of evaluation manpower. Since, in reality, most evaluation is carried out systematically and comprehensively by organisations rather than by individual evaluator, the structures of evaluation unit may affect the utilisation of evaluation manpower. Evaluators must also possess expertise and experience in evaluation itself. In order to evaluate the quality of evaluation manpower, we must analyse whether it has sufficient knowledge and techniques in relation to policy evaluation, a high expertise with respect to evaluation objects and evaluation experience in the sectors concerned.

Therefore, in order to evaluate the quantitative factors of evaluation manpower, meta-evaluators must examine the total number of evaluation manpower, the ratio between internal and external manpower, the availability of the external expert groups and so on. Also, to evaluate the quality of evaluation manpower, we need to comprehensively examine educational backgrounds of manpower, the length of their service to the evaluation work, their major interest and so on.

## **(2) Evaluation Unit**

To perform the policy evaluation actually, it requires evaluation units which can systematically arrange and manage the evaluation activities. The evaluation of evaluation units, in particular, is needed more in cases where policy evaluations are carried out by special institutions who have official authority, and not in the case of individual evaluators. With respect to this, it is necessary to evaluate the availability of separate and special evaluation unit, and its independence from other units in the internal organisational hierarchy. For effective and efficient policy evaluation, it is very important to establish special evaluation units and sometimes ones that are independent in their functions.

In general, it is said that an evaluation unit has the following three functions : (1) to organise and manage the overall evaluation process in accordance with the related guidelines, (2) to actually execute the evaluation in accordance with its purposes and (3) to carry out the evaluation research to develop evaluation models, standards, methods and so on and apply them to the actual evaluation stages (Lee et al., 1994a). Therefore, in

order to evaluate evaluation units, it is necessary first to examine whether there is a special and separate evaluation unit available. Next, it is necessary to analyse whether the evaluation unit is carrying out such roles as evaluation organiser, evaluation performer and evaluation researcher.

### **(3) Evaluation Information**

The evaluation unit or evaluator should have sufficient information relating to the evaluation objects. The usefulness of evaluation information will be determined by the quantity and quality of information sources. While the quantity of information is related to whether evaluators can gather and utilise sufficient amounts of information, the quality of information is related to whether such information is reliable to. For this reason, in order to assess the evaluation information, the meta-evaluator has to firstly investigate how many information sources are available to evaluators and then to judge the sufficiency and reliability of the information obtained from them.

#### ***3.4.2.3. Evaluation Performance***

Evaluation performance deals with the substantial elements during the evaluation. As already indicated, since the contents of evaluation performance, including the evaluator, evaluation method, and evaluation criteria and indicators, have been recognised as the core factor of meta-evaluation, there is a lot of research on this subject.

#### **(1) Evaluator : Internal or External Evaluator**

It is not right to confine evaluators to a certain social class or interest group, because everyone in democratic countries can evaluate all government policies and programmes in a direct or indirect way. However, even though anyone can be an evaluator, it is important to distinguish external evaluators from internal evaluators in the light of the evaluation purposes (Glaser and Backer, 1972 : 56-60). This is a quite complicated matter including decisions about who is a better evaluator and whose interests should be most emphasised : those of the evaluation commissioner, the evaluation planners, the evaluators or the evaluation clients. In addition, just as internal and external evaluators are often more complementary rather than exclusive each other, so it is desirable to select the evaluators in relation to the evaluation purposes (Vedung, 1997 : 170-174). Consequently, in evaluating the evaluator, the meta-evaluator should judge whether the

evaluator who can best achieve the evaluation purposes has been selected or not.

## **(2) Evaluation Method**

It is necessary to adopt suitable evaluation methods which can cover evaluation design and the collection of data concerned and its analysis. In general, both scientific methods and non-scientific methods can be used for policy evaluation. Scientific evaluation means that evaluator adopts scientific, systematic and logical tools in evaluation. It is classified into qualitative and quantitative methods, according to the methodology of data analysis (Patton, 1990 ; Rossi and Freeman, 1993 ; Wholey et al., 1994). It is also classified into experimental methods, quasi-experimental methods and non-experimental methods, according to the strictness of the research design (Campbell and Stanley, 1966). Non-scientific evaluation covers the evaluator's subjective judgment and the evaluations for certain political purposes.

However, in policy evaluation, scientific and systematic evaluation methods are more meaningful than non-scientific ones. Since evaluation methods may vary according to the evaluation objects, evaluation timing and evaluation criteria applied, it is necessary to use various of them in actual evaluation (Filstead, 1979 : 42). In addition, since evaluators cannot use all the available evaluation methods due to limited resources, they are always faced with the problem of selecting the evaluation methods which are most suitable to the evaluation purposes (Alemi, 1988 : 765). Hence, in order to evaluate evaluation methods, the meta-evaluators should examine whether those that have been adopted are suitable, given the resource constraints, after that, also judge their effectiveness.

## **(3) Evaluation Criteria and Indicators**

When we intend to evaluate certain programmes or policies, we normally judge their success or failure on the basis of certain criteria or indicators. Therefore, the selection of criteria or indicators greatly affects the validity and reliability of evaluation findings. On this subject, it is certain knowledge that, while some researchers simply present the general evaluation criteria, others propose more detailed and suitable ones depending on the policy components and policy/programme types (Vedung, 1997 : 338-358). From this discussion, it is apparent that, in order to carry out the reliable and appropriate evaluation, different evaluation criteria should be adopted for the evaluation content on the one hand and evaluation procedure on the other. Therefore, in the evaluation of evaluation criteria,

the meta-evaluators should examine whether both substantial and procedural criteria have been adopted, and whether suitable evaluation criteria for each different policy and programme have been applied.

Meanwhile, since each evaluation criterion can be measured by the individual evaluation indicator during the actual evaluation process, evaluation indicators should be selected by taking evaluation objects and evaluation timing into consideration. Regarding the evaluation indicators, there are discussions about the combinations of quantitative and qualitative indicators to be applied in the evaluation process. While some researchers are used to differentiating them in the actual evaluation process (Chung and Grupp, 1990), others are not (Geisler, 1994). Since quantitative indicators help evaluators to compare the outputs or outcomes of programmes/policies evaluated with those of another, they make it easier to maintain the objectivity of evaluation findings. However, only adopting quantitative indicators makes it difficult to qualify policy outputs and outcomes, and to measure the policy impacts on society and the economy exactly. Hence, if we intend to perform impact evaluations, which are more complicated and abstract than output evaluations, it is absolutely necessary to adopt a combination of both quantitative and qualitative indicators. Consequently, in evaluating evaluation indicators, both the combinations between quantitative and qualitative indicators and the suitability of each indicator applied should be examined.

#### ***3.4.2.4. Evaluation Utilisation***

The evaluation utilisation deals with the feedback process whereby the findings of a primary evaluation will affect the next stages of the policy process. While Larson and Berliner (1983: 159-160) confine evaluation utilisation to the detailed decisions after evaluation, it is desirable that evaluation utilisation should include all contents which the evaluation researchers have said should fall under this topic. According to existing research, evaluation utilisation should include both the instrumental utilisation (Alkin, Daillak and White, 1979 : 225-226) and the conceptual utilisation of evaluation findings (Rich, 1977 : 200 ; Luukkonen and Ståhle, 1990 : 361-363 ; Rossi and Freeman, 1993 : 443 ; Vedung, 1997 : 375-376). I shall now discuss the production and dissemination of the evaluation report which are pre-conditions for evaluation utilisation, and to examine the types of evaluation utilisation.

### **(1) Evaluation Report : Production and Dissemination**

Just as the final findings of evaluations are embodied in the evaluation report, so its production and disseminations are essential factors in evaluation utilisation (Horvat, 1999 : 538). The evaluation report should be prepared in such a way as to help stakeholders to make an attention to the contents. Regarding the contents of the evaluation report, meta-evaluators should examine whether the evaluation purposes and objects, evaluation methods, data analysis and conclusions are correctly described (Hoogerwerf, 1992 : 220-221).

Since one of the evaluation purposes is to promote the rationality of the policy process by producing useful information, the evaluation findings should be disseminated to all stakeholders. If the evaluation findings once successfully reached are disseminated to only specific institutions or clients, similar institutions may carry out the same evaluation work, or the audiences concerned cannot be helped by earlier evaluations. Therefore, meta-evaluators should investigate whether evaluation reports are disseminated on time to all stakeholders involved in the policies and programme evaluated, and whether these are delivered to the key stakeholders through the special reporting channels (Geroghiou, 1995 : 185).

### **(2) Utilisation Type : Instrumental and Conceptual Utilisation**

For the evaluation findings to be more important references in following policy processes, they should be widely utilised from the conceptual perspectives as well as the instrumental ones.

Normally, the evaluation findings result in the expansion, termination or maintenance of the policies and programmes evaluated. This is one kind of instrumental utilisation in evaluation findings. Just as the most direct utilisation of evaluation leads to budget increases or decreases depending on the evaluation findings, it is necessary for policy evaluations and budget allocations to be closely linked. Therefore, meta-evaluators should examine whether evaluation processes and budget processes are systematically integrated, and whether the evaluation findings result in increases or decreases in budget. Meanwhile, for the evaluation findings to be directly reflected in policies and programme, the evaluation organisations must have the mission to make a corrections based on the evaluation findings. Accordingly, in the evaluation of the instrumental utilisation, the



relationships between evaluation findings and budget allocations, and the availability and types of corrections by the evaluation organisation should be examined.

As well as short-term and direct (instrumental) utilisation, long-term and indirect (conceptual) utilisation should be emphasised. Therefore, meta-evaluators should examine the extent to which the evaluation findings will contribute to the future development of the pertinent policy or programme from long-term and strategic perspective.

### 3. 5. Design of Research Framework : Meta-evaluation Model

When meta-evaluation is actually performed in relation to an individual policy evaluation, its detailed components may vary depending on the policies or programmes evaluated. Also the importance and priority of each of them may not be same for all policies and programmes evaluation. However, despite the flexibility and changeability of the components, in order to carry out effective and reliable meta-evaluation, some core components should be considered. In this respect, the meta-evaluation model that I intend to propose as a research framework for my thesis comprises the four main components and 11 sub components shown in <Table 3-6>.

<Table 3-6> Research Framework : Meta-evaluation Model

Main Components of Meta-evaluation	Sub Components of Meta-evaluation
Evaluation Paradigm	Evaluation purpose
	Evaluation object : level and scope
	Evaluation type
Evaluation Resources	Evaluation manpower : quantity and quality
	Evaluation unit
	Evaluation information
Evaluation Performance	Evaluator : internal or external
	Evaluation method
	Evaluation criteria and indicator
Evaluation Utilisation	Evaluation report : production and dissemination
	Utilisation type : instrumental and conceptual utilisation

Using this meta-evaluation model, I shall first analyse and judge the effectiveness and appropriateness of the R&D evaluation systems and their exercises in both the UK and Korea. Secondly, I shall compare the two countries' R&D evaluation systems and investigate their strengths and limitations. After that, on the basis of these meta-evaluations and comparisons, I shall discuss how the UK evaluation systems in the S&T and R&D sector could be applied to the Korean cases, whose background and current administrative situations are different from those of the UK.

# **CHAPTER 4. CASE STUDY OF THE UK NATIONAL R&D EVALUATION SYSTEM**

## **4. 1. Introduction**

As already discussed, the UK has the most advanced national R&D evaluation system of all developed countries and also has a relatively long history of active exercises in this area. In this chapter, I conducted the case study of the UK R&D evaluation system at the national level. This chapter comprises two main sections excluding introduction : Section 4.2 relates to the current state of the R&D evaluation systems, Section 4.3 deals with the recent evaluation exercises in the S&T and R&D sectors. More detailed, Section 4.2 includes broad and, where available, specific, information which describes both the past and current state of the UK R&D evaluation systems. It also covers more detailed issues such as government SET organisations, a brief history of the development of R&D evaluation, the current hierarchy of national R&D evaluation systems, and major guidance and consultation documents. Section 4.3 includes recent five evaluation exercises covering the evaluation of science and research policy, the evaluation of research institutions and the evaluation of S&T programme and R&D projects.

For the work in the second and third section, I firstly reviewed the existing research articles and reports including evaluation report, as well as many documents published or posted on the website of the organisations concerned. I also conducted interviews with previous and current staff members of the related government organisations concerned, in order to complement what was unclear or unconfirmed in the literature review. Interviews were mainly conducted face-to-face and by e-mail, depending on the specific interview situations.

## **4. 2. Current State of the R&D Evaluation Systems in the UK**

In this section, in order to provide a general understanding of Science, Engineering and Technology (SET) activities in the UK as a whole, I shall first describe the government

organisations in the SET sector. Next, I shall provide a brief history of the development and an account of its current state of the national R&D evaluation system in the UK, with particular emphasis on the organisational hierarchy between them and their guidance and consultation documents of both government departments and the research councils (in particular, the BBSRC).

#### ***4.2.1. Government Organisations for SET***

Regarding this subject, it seems that there are a lot of different approaches to deal with. However, following the overall UK governance system in the SET sector, I intend to suggest the simplified typology of the government organisations, in accordance with their main roles and missions in the policy process for SET, even though they are sometimes slightly overlapped or duplicated with other organisations.

##### ***4.2.1.1. Overview of the Governance System of the SET***

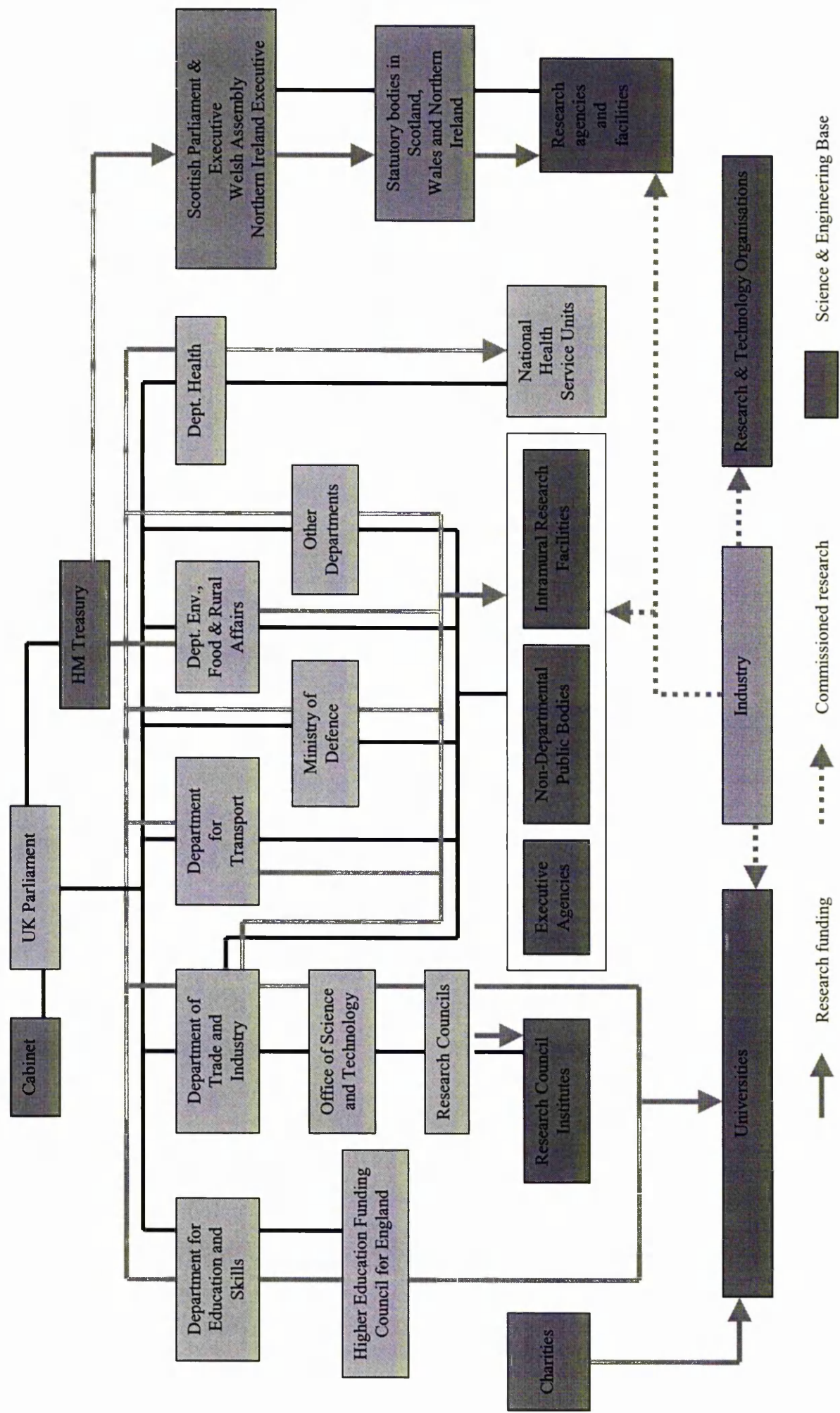
Over the last some decades, the UK has witnessed a number of changes in the organisation of its SET structures and policy-making mechanisms. In particular, it has been argued that the creation of the OST within the Cabinet Office in 1992 and its transfer to the DTI in 1995 was one of the turning points in the SET sectors. This meant that the control powers of the central government to the individual government departments continued to increase and that the close connection between the SET activities and the industrial competitiveness became to be more emphasised. The process of centralisation continued, under the Labour Government which came to power in May 1997, with the merger of the some government departments and the creation of a number of committees for the advisory and policy-making at the central government level (Cunningham, 1998b : 69).

However, the emphasis will place on not the history of its development and extension but an outline understanding of the governance system of the SET as a general, within the limit of the providing information of the policy process,<sup>1</sup> in particular, of the policy

---

<sup>1</sup> Regarding this subject, there are many references (Jones et al., 2001 ; Hills, 1998 ; Parsons, 1997 ; Ham and Hill, 1993 ; Hogwood and Gunn, 1984). On the basis of existing discussions, in my thesis, I will use the term policy process to refer to a series of procedures covering a problem definition, an agenda setting, a

formulation and policy evaluation. Therefore, I shall mainly provide the overall organisational structure of government SET activity, particular in terms of the responsibilities for policy-making, the S&T and R&D responsibility and the flow and level of research funding. The principal UK government components in the SET sectors including devolved executive organisations in the Scotland, the Wales and the Northern Ireland are shown in (Figure 4-1).



(Figure 4-1) Overview of the UK System of Governance of SET

Meanwhile, in the UK, SET activities are categorised into two types. One is the basic and strategic SET for future national competitiveness and economic development as a whole, which is mainly funded by the DTI/theOST and the DfES. The other type is mission-oriented SET for policy purposes which has been funded by civil departments such as the DEFRA, the DfT, the DoH and so on (Hills and Dale, 1995 : 36-37). There are slight differences between the policy process of the two types of SET activities.

Having given this background, I shall give a brief description of government organisations and their roles in the area of SET in the UK. Government organisations involved in SET could be broadly grouped into those whose purpose is policy formulation, those who deal with budget allocation and major R&D performers, according to their major roles in the SET policy process. These organisations are summarised in <Table 4-1>. As already indicated in Chapter 1, the research scope and level of my thesis are focused on the evaluation systems and exercises which HM Treasury, the DTI, the OST and research council (in particular, the BBSRC) have been operating, so my exploration of this subject will be mainly devoted to organisations related to the area of basic and strategic SET activities.

**<Table 4-1> Typology of Government Organisations for SET in the UK**

<div>Policy Process</div> <div>SET Type</div>	Policy Formulation		Budget Allocation	Major R&D Performers
	Central Govern. Level	Department Level		
<b>Basic &amp; Strategic SET</b>	- HM Treasury - Cabinet Office	- DTI/OST - DfES	- OST/RCs - HEFCs	- RCs Institutes - HEIs
<b>Mission-Oriented SET</b>	- OST	- DEFRA, DfT, DoH, MoD etc.		- PSREs

\* This is schematic for analytical purposes. In reality, the discussions between basic and mission-oriented SET are blurred.

#### ***4.2.1.2. Organisation for Policy Formulation***

The organisations involved in the formulation of policy for SET are again divided into two types : one includes “central government organisations”, such as HM Treasury, the Cabinet Office and the OST, and the other includes “department organisations”, such as

the DTI, the DfES, the DEFRA, the DfT, the DoH and so on. The central government organisations set up a general and basic framework for S&T policy through various mechanisms, including budget allocation and the issue of guidelines that affect basic and strategic SET as well as mission-oriented SET, but do not prescribe in detail how it is formulated and implemented. The policy process at the more detailed level, in particular policy implementation, is almost entirely devolved to the individual departments directly concerned. The department organisations have a mission to formulate more detailed and specific SET policy, for example, articulating their needs for SET, formulating programmes and projects and managing the research process as customers in their domain, in line with the guidelines or advice of the central government organisations.

### **(1) Central Government Organisations**

**HM Treasury** : It mainly exercises influence through three-yearly Comprehensive Spending Reviews (CSR) and their annual updates. These determine the amount of money available to departments in total and in broad categories. The budget allocation that each department secures for SET is clearly likely to reflect wider national priorities. The allocations are supplemented by Public Service Agreements (PSAs) between HM Treasury and each department, which consist of a negotiated list of objectives that the department pursues in return for the money it receives. Each PSA is underpinned by a Service Delivery Agreement (SDA) which fills in some, but not much, detail about the way in which the department seeks to achieve its objectives. In a broad sense, this process sets a framework for SET activities, but rarely in any specific way (Hills, 2001 : 64).

In the SET sector, in particular, HM Treasury has conducted a “cross-cutting review of science and research”, one of the cross-cutting reviews conducted in many areas, that will contribute to the Spending Review (SR) for the coming year. This includes a review of science funding by the DTI/the OST and the DfES in the UK SET base and the effectiveness of civil science and research programmes commissioned by devolved government departments such as the DEFRA, the DfT, the DoH and so on, in order to ensure that they deliver maximum long term benefits to the economy and the quality of life (HM Treasury, 2000a ; 2001a). Also, HM Treasury has issued many white papers and guidelines, including *The Green Book : Appraisal and Evaluation in the Central Government (1997)*, in order to improve the policy formulation and evaluation in the areas under its responsibility which includes SET. While the HM Treasury’s



documentation sometimes contains very detailed and open technical guidance on certain aspects of the policy evaluation, its advice and guidelines are formulated in fairly general terms and are very much in line with practice in the devolved departments, for example, as like the DTI's ROAME statement.

**Office of Science and Technology** : The OST was established in 1992 within the Cabinet Office. Since 1995, it has been part of the DTI where it provides the central focus for the development of government policy on science and technology nationally, across the broad spectrum of government and internationally. It has responsibility for the allocation of the science budget for the research councils and universities, the work of the research councils and for implementing the policies set out in the 1993 and 2000 science and innovation White Papers (British Council, 1999 : 13).

However, the OST does not intervene extensively in the detail of government departments SET activities, but confines itself to influencing strategy and the key principles applying to the development and presentation of scientific advice for policy making. The first guidance on these key principles was issued in 1997 as *The Use of Scientific Advice in Policy Making : A Note by the Chief Scientific Advisor*. Two reports have been published on the implementation of this document, which was seen as not quite sufficiently clearly focused. This led to revised guidance, namely *Guidelines 2000 : Scientific Advice and Policy Making*, published in July 2000, which stresses the importance of identifying issues on which scientific advice or research may be needed as early as possible, acquiring the best expertise, asking the right questions and ensuring transparency. This reflects the emphasis on what is, essentially, peer review, although it is mediated in a variety of forms (Hills, 2001 : 64).

**Cabinet Office** : It has recently taken significant steps to improve policy making across all areas of government. These started with the *White Paper : Modernising Government*, published in March 1999, and were followed by more detailed documents. Of these, the most significant ones in the present context are *Professional Policy Making for the Twenty First Century (1999)* and, to a lesser extent, *Adding It Up (2000)*. These emphasise the importance of evidence-based policy making and evaluation. They do not go into any detail on methodology. The Cabinet Office, however, intends to set up a centre of excellence on evaluation within its Performance and Innovation Unit. It does not

envisage this centre undertaking specific evaluations, but sees it as encouraging best practice and advising departments when appropriate. While it is not planned that it should have prescriptive power, it is likely to be influential. Departments that reject its advice will need to have good arguments for having done so when they seek funds from HM Treasury to continue the policies concerned (Hills, 2001 : 65).

In summary, the central organisations wish to ensure that, wherever appropriate, SET policy in both the basic and strategic field, and also the mission-oriented field is underpinned by sound evidence. They are not much concerned in this context about whether that underpinning derives from the existing knowledge and expertise of advisers to whom they have access or from specially commissioned research. While these organisations influence and offer guidance on the broad priorities, strategy and general principles, they do not involve themselves, on a day-to-day basis, in the detail of the executive departments that handle their scientific arrangements (Hills, 2001 : 65). These generally decentralising tendencies are complemented by measures to improve interdepartmental co-operation on cross-cutting policies. They do not, however, involve central organisation intervention except as a last resort (Cabinet Office, 2000b).

## **(2) Department Organisations**

Regarding the organisations for policy formulation at the level of the individual department, the DTI and the DfES have played a major role in basic and strategic SET activities as a whole.

**Department of Trade and Industry** : The DTI, excluding the OST, is the department responsible for promoting competitiveness, enterprise and innovation in industry at the national level, thus contributing to economic development and the quality of life. The DTI funds around £300 million of research annually and is by far the most important civil department in terms of civil R&D expenditure on science. Much of this is applied research in support of industry. Some research is to maintain national measurement standards, and some is in pursuit of the national space programme, which is carried out largely through the European Space Agency. Almost all of the DTI's research programmes are contracted out. Also, the DTI is responsible, through the OST, for the science budget, the research councils and cross-departmental co-ordination and the development of government policy on SET (Cunningham, 1998b : 86 ; Hills , 2001 : 67).

As a result, the DTI has two separate missions, which are very close but a little different. These are to promote competitiveness and innovation in industry through applied research programmes, and to contribute to economic development and quality of life through science and basic technology programmes.

**Department for Education and Skills** : Another department responsible for the basic and strategic SET activities is the DfES. Its two broad aims are to raise the level of educational achievement and skills of the population and to promote an efficient and flexible labour market. It has overall responsibility for education and training matters at all levels, from schools to higher education and vocational training. The department carries out policy formulation in the area of basic and strategic SET, in the course of its responsibility for the research programmes in the higher education sector. Hence, much of the scientific expertise on which the department draws comes from academia and much of the research that it commissions contributes both to policy formulation and to the science base generally (Cunningham, 1998b : 114 ; Hills, 2001 : 63).

#### ***4.2.1.3. Organisation for Budget Allocation***

In the area of basic and strategic SET, there are two clusters of organisations for budget allocation, as already indicated. One comprises the OST/the research councils, and the other comprises the Higher Education Funding Councils (HEFCs) for England, Scotland, Wales and Northern Ireland. In the UK, a dual support system of higher education funding has been running, in order to fund research activities in the universities.

##### **(1) The OST and the Research Councils**

The UK research councils were established by the Royal Charter as independent Non-Departmental Public Bodies (NDPBs), in order to support research, postgraduate training and the public understanding of science. They receive most of their funding, which now approaches some £2,000 million each year, via the OST. There are six grant awarding councils and one that manages research facilities for other research suppliers (DTI, 2001b : 4). These are as follows :

- The Biotechnology and Biological Sciences Research Council (BBSRC)
- The Engineering and Physical Sciences Research Council (EPSRC)
- The Economic and Social Research Council (ESRC)

- The Medical Research Council (MRC)
- The Natural Environment Research council (NERC)
- The Particle Physics and Astronomy Research Council (PPARC)
- The Council for the Central Laboratory of the Research Councils (CCLRC)

The Science Budget, which both the DTI and the OST are responsible for, is allocated to the R&D performers via two stages. It is allocated to each research council by the OST. Then each individual research councils allocate it to both the research council's institutes and universities. The allocation of funding between the research councils is agreed as part of each spending review and reflects both the current relative priorities attributed to each and their established level of expenditure. The OST has established a top-level management forum, the Chief Executives of the Research Councils (CERCs), chaired by the Director General of Research Councils (DGRC), which enables strategic issues such as funding allocations to be addressed outside the parochial concerns of the individual councils. The research councils individually fund research programmes in universities and in institutes they own or support, and other research suppliers with varying degrees of priority. The research programmes are either "responsive" or "targeted", and all rely on peer review to confirm that research proposal's merit the funding requested. All of the councils have made considerable use of "initiatives" to generate research programmes in new areas of science or of a specific format, for example, teamed with industry or other partners. Many of these initiatives have operated at the boundaries between research councils (DTI/OST, 2001b : 5).

The research councils support research in the higher education sector directly through the provision of research grants, fellowships and postgraduate student support. Depending on the research council in question, these grants can support research projects, programmes or designated centres of research activity. Some research councils provide indirect support through the provision of large-scale facilities. Also, a number of research councils, which have their own research institutes or research units, directly support research programmes or projects. Despite some variation, in most cases the majority of research councils' expenditure is on 'basic-oriented' and 'applied-strategic' research. The notable exception to this is the PPARC, where expenditure is almost entirely devoted to 'basic-pure' research (Cunningham, 1998c : 178-179).

## **(2) Higher Education Funding Councils (HEFCs)**

The Higher Education Funding Councils (HEFCs) for England, Scotland and Wales were founded in 1992 under the terms of the Further and Higher Education Act 1992 and the Further and Higher Education (Scotland) Act 1992. They are non-departmental public bodies (NDPBs) appointed and responsible to the Secretary of State for Education, the Secretary of State for Scotland and the Secretary of State for Wales, respectively. Meanwhile, the Northern Ireland Higher Education Council (NIHEC) provides advice to the Department of Education for Northern Ireland (DENI), based on advice from HEFC for England (Cunningham, 1998c : 174).

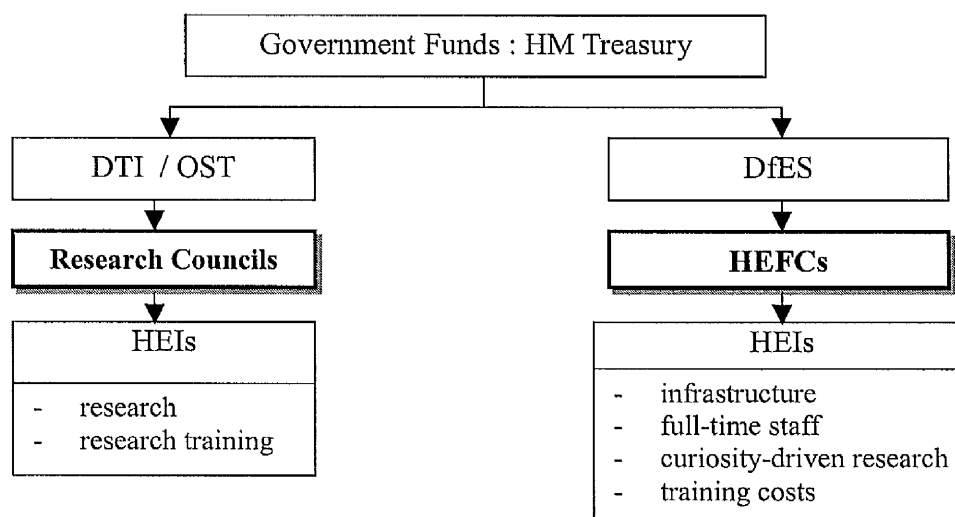
Each HEFC allocates research funds, in the form of a block grant, to the universities within its area of responsibility, in order to support teaching, research and related activities. In the case of the SHEFC, the HEFCW and the DENI, regional initiatives may be employed to respond to local circumstances. The split between teaching and research is decided annually, although the division is generally consistent from year to year. The majority of these funds are allocated according to set formulae, which consider the size and activities of individual universities and the quality of their research. Whilst the priorities of the individual HEFCs may vary somewhat, the research funding criteria reflect the quality and volume of research work and the levels of activity supported by other funders and users of research. Individual universities are then free to distribute the money internally according to their own judgement, as long as the funds are used for the purposes for which they were provided. However, a further level of accountability has been created in that universities are now expected to take account of Technology Foresight priorities (Cunningham, 1998c : 174-175).

In order to allocate research funds to the higher education sector, HEFCs have carried out a Research Assessment Exercise (RAE) every five years. The purpose of the RAE is to inform the higher education funding bodies in their allocation of 'quality-related research' (QR) grants to the institutions that they fund. From 1996, the RAE has been mounted jointly by all four HEFCs. The information requested from and supplied by higher education institutions (HEIs) provides the basis for peer review assessment of research quality by a number of specialist panels. These assessments are carried out on the basis of standardised written submissions, which include both quantitative and descriptive elements. HEIs are also free to bring to the attention of the panels any other

matters which they consider relevant to the judgement of research quality (Cunningham, 1998c : 176-177). As a recent exercise, between spring 1999 and December 2001, RAE 2001, the fifth exercise, was conducted to aim at allocating approximately £5 billion of research funding to the HEIs. In the RAE 2001, there were 12 categories of information in the submission : (1) summaries of all academic staff, (2) details of research-active staff, (3) research support staff and research assistants, (4) up to four items of research output for each researcher, (5) information about the research environment, structure and policies, (6) strategies for research development, (7) qualitative information on research performance and measures of esteem, (8) amount and sources of research funding, (9) numbers of research students, (10) number and sources of research studentship, (11) numbers of research degrees of awarded and (12) indicator of peer esteem (RAE, 2001 : 4).

### (3) The Dual Support System of Higher Education Funding

As already indicated, higher education funding in the UK is based on principles of competition, selectivity, accountability and a plurality of funding sources : the Higher Education Funding Councils and the Research Councils. HEFCs' block grants cover research and teaching activities, and universities are accountable for the way in which they allocate their research funds. HEFCs' funds provide the universities with underpinning for basic curiosity-driven and strategic research, and for longer-term collaborative work. In this, they support research training and provide general support, including research libraries, academic staff, central computing and premises for research council projects.



(Figure 4-2) Overview of the Dual Support System of HEIs

The research councils are responsible for the provision of specific funding for research programmes and projects. The grants they make as part of the dual support system cover the direct costs associated with projects and include a contribution towards indirect costs, including a 40% addition to the direct staff costs (Cunningham, 1998c : 172-174). The UK's dual support system for higher education institutions is shown briefly in (Figure 4-2).

#### ***4.2.1.4. Organisation for R&D Performance***

In the UK science base, research institutes under the responsibility of the research councils and HEIs have played a major role in carrying out the national R&D programmes and projects. Many parts of their research grants normally come from the research councils and the HEFCs. In addition, they may also win research funds from the related government departments (e.g. the DTI, the DEFRA, the DfT and the DoH etc.), industry and the European Union.

##### **(1) Research Councils Institutes**

In order to carry out long-term strategic, often multidisciplinary research, some research councils sponsor their own institutes and central facilities. Also, some support a lot of research centres and research units in partnership with universities and other non-profit research funding organisations such as the Wellcome Trust (DTI/OST, 2001b : 5). There is a huge range in the size of the RCs' institutes. By far the largest RCs' Institute is the Roslin Institute under the BBSRC, which has over 600 employees. At the other end of the scale, several research units run by the MRC look into a specific aspect of medical science and have fewer than 20 persons. Also, each research councils institutes have a various types of organisational status. While institutes and research centres of the BBSRC, the NERC and the PPARC are separate legal entities, some research councils' centres and units have been established in partnership with or attached to HEIs (HM Treasury, 1999 : 10 ; DTI, 2001b). As of August 2001, the number of research institutes, centres and units that individual research councils own or support is summarised in <Table 4-2>.

**<Table 4-2> Number of UK Research Councils Institutes**

RC Type	BBSRC	ESRC	EPSRC	MRC	NERC	PPARC	CCLRC	Total
Independent Institutes/ Centres	8	-	-	3	7	3	-	21
Research Centres with HEIs	6	20	15	3	-	-	-	44
Research Units/Groups within HEIs	6	-	-	32	7	-	-	45
Large Facilities/ Resource Centres	-	10	-	-	-	-	3	13
Other	-	-	1	3	-	-	-	4
<b>Total</b>	<b>22</b>	<b>30</b>	<b>16</b>	<b>41</b>	<b>14</b>	<b>3</b>	<b>3</b>	<b>129</b>

Source : DTI (2001b).

## **(2) Higher Education Institutions**

The Higher Education Institutions (HEIs) are the largest performers of basic and long-term strategic research in science and technology, including medicine and environment in the UK. For example, in 1999-2000, of a total government net expenditure on R&D of £6,529 million, 21% (£1,338.5 million) was allocated to the research councils and 18% (£1,157.1 million) was to Higher Education Funding Councils. In broader terms, R&D performed in the higher education sector accounted for just under 20% of civil gross domestic expenditure on R&D (GERD) and 0.4% of GDP (OST, 2001c). Moreover, in recent years, HEIs have also become responsible for increasing the amount of short-term, applied and near-market research, generally commissioned by or in collaboration with industry, government and other bodies. For example, as at 1997, industry spent around £167 million on research in universities (Cunningham, 1998c : 159). As of December 2001, there were 156 universities and other HEIs in the UK such as 122 HEIs in England, 18 ones in Scotland, 13 ones Wales and 3 ones in Northern Ireland (RAE, 2001b).

### ***4.2.2. Brief History of the Development of R&D Evaluation***

It has been argued that the R&D evaluation system is one of the particular types of policy



evaluation. Even though there are some differences between policy evaluation and R&D evaluation in terms of, in particular, evaluation methods and evaluation indicators and so on, the two evaluation systems have rather similar and sometimes the same contents and processes. So, in this part, I shall give an explanation of the history of the development of the R&D evaluation system in the UK, in conjunction with that of the policy evaluation system in general.

#### ***4.2.2.1. Development of Policy Evaluation***

In the UK, evaluation activities in the public sector have been executed in a variety of forms for many years. By 1914, the Exchequer and Audit Department, which is now the National Audit Office (NAO), employed 250 staff, some of whose work was similar to current evaluation activities. Nowadays, the work of the NAO serves some of the same purposes, being designed to make government accountable through Parliament to the public for using taxpayers' money efficiently, in order to prevent waste or fraud. It also reports on the effectiveness of arrangements for securing value for money in public spending. However, only incidentally, it is expected to contribute positively to the formulation or execution of public policy.

Evaluation as it is now understood only began to emerge when other pressures made government more aware of the demands of efficiency for its own purpose rather than that of taxpayers. Perhaps the earliest obvious manifestation of this was the system of Programme Analysis and Review (PAR). Instituted by the 1970-1974 Conservative Government, PAR had some of the characteristics of evaluation. At about the same time, the DTI started to review its activities on a rolling basis. This amounted to a rather non-rigorous form of evaluation, but the arrangement did not continue after the Department was divided into separate parts in 1974. These early developments were relatively unsystematic, and in some cases superficial, compared to the burgeoning of interest in policy evaluation since about 1980. More specifically, greater emphasis on evaluation has stemmed partly from exogenous intellectual developments derived, for example, from economic and management science. In addition, within government, the primary impetus has come from increased pressure on budgets forcing often-draconian prioritisation. Finally, without the enormously increased processing capacity and accessibility of computers, much evaluation work that has occurred would have been

effectively impossible (Hills and Dale, 1995 : 37).

An early indication of the new enthusiasm for evaluation was the Financial Management Initiative (FMI), which was launched by HM Treasury in 1982. This was intended to promote in each department an organisation and a system in which managers at all levels have a clear view of their objectives and a means to assess and, wherever possible, to measure output or performance in relation to them (HM Treasury, 1983). This was followed by a series of further guidelines and related White Papers. Of these, in particular, it has been said that *Policy Evaluation : A Guide for Managers*, issued by HM Treasury in 1988, was a general and comprehensive guide to policy evaluation. This guidance is addressed primarily to managers with line responsibility for policies and programmes and projects, rather than to specialist advisers. As a result, its purpose is to help such managers understand the main concepts and to plan and mount evaluations in areas for which they have a responsibility (HM Treasury, 1988 : 2). It also contains a section devoted exclusively to the evaluation of science and technology programmes, along with sections on a range of other policy areas (HM Treasury, 1988 : 17).

Subsequently, HM Treasury has issued *Economic Appraisal in Central Government : A Technical Guide for Government Departments* in 1991. This guide is intended to help officials and managers in central government to appraise and evaluate expenditure proposals effectively. It describes the principles of appraisal and project evaluation, and provides specialist readers with technical guidance on issues of wide application in central government including agencies. For many applications, departments wish to develop and maintain their own guides, tailored to their specific needs and consistent with the conventions set out in HM Treasury's. The previous application of the guide has been in the appraisal of the capital projects. However, this guide has been extended to cover a wider range of expenditure applications to which same general principles apply, and to give more prominence than before to post-project evaluation (HM Treasury, 1991 : 5).

Following the previous guidelines for policy evaluation, HM Treasury issued "*Appraisal and Evaluation in Central Government*" again in 1997. This publication is known as "*The Green Book*". It contains very detailed and often technical guidance on *ex ante* appraisal and *ex post* evaluation. It covers such techniques as cost benefit analysis, including handling risk and uncertainty, the cost of capital, the discounting of rates of

return and so on. It also discusses the managing of evaluation and provides a checklist of necessary action. In practice, however, much of the detailed advice is only applicable to very large undertakings such as building a hospital or a major road. It would be too cumbersome for any but exceptionally large research projects. The advice on managing an evaluation is in fairly general terms and very much in line with practice in, for example, the DTI (HM Treasury, 1997). Meanwhile, HM Treasury published a new edition of the *Green Book* in January 2003, which will take full effect from 1 April 2003 (UKES, 2001 ; HM Treasury, 2003).

Recently, another important guideline has been established for all government department policy making, including that related to policy evaluation. The Cabinet Office published the *White Paper : Modernising Government* in 1999. This is central to the government's programme of renewal and reform, and is also a long-term programme of improvement for the public services. The white paper set out five key commitments : forward looking policy making, responsive public services, quality public services, information age government and valuing public service. Regarding policy evaluation, in line with the commitment to quality public services, the government has to undertake to review all central and local government department services and activities over the next five years to identify the best supplier in each case. It must also undertake to monitor performance closely so that it will strike the right balance between intervening where services are failing and giving successful organisations the freedom to manage (Cabinet Office, 1999a : 6-7). Under the general principle of this white paper, the Cabinet Office also issued *Guidance on Agency and NDPB Quinquennial Review : How to Review Agencies and Non-Departmental Public Bodies (NDPBs) to Improve the Quality and Effectiveness of Public Services* in February 2000. This new guidance contains detailed processes and methods which are seen as necessary for reviewing the organisational status and performance of agencies and NDPBs (Cabinet Office, 2000a).

#### **4.2.2.2. Development of R&D Evaluation**

Turning from the above general developments in policy evaluation to those specifically affecting R&D, the DTI had set up a small Science and Technology Assessment Unit (STAU) in the early 1980s, partly in response to the general HM Treasury pressures mentioned above. In 1984, the DTI greatly expanded its Assessment Unit and formally

instituted evaluation across the whole range of its programmes. Other departments adopted approaches suited to their particular circumstances (Hills and Dale, 1995 : 38).

In 1983, the DTI, along with the MoD and the Science and Engineering Research Council (SERC), instituted the continuous evaluation of the major information technology research programme known as the Alvey Programme. This evaluation, completed in 1991, was carried out by PREST at the University of Manchester and SPRU at the University of Sussex (Guy and Geroghiou et al., 1991). It has been almost certainly the largest single evaluation undertaken so far, and, even more so, if taken together with its successor, the Information Engineering Advanced Technology Programme, which was carried out by PREST alone, and completed in 1995 (Cameron et al., 1996). Through the execution of these evaluations, the DTI was able to experiment with a mixture of internal and external evaluation approaches and with real-time as well as *ex post* evaluations during the later 1980s. This was also a time when the DTI began to plan programmes more formally through the use of so-called ROAME statements. All proposed programmatic interventions by the department were to be submitted using a standard format intended to ensure that the case for intervention was explicit and agreed, and that the programme would be adequately managed. The ROAME statement comprises five sections : Rationale, Objectives, Appraisal, Monitoring and Evaluation (Guy and Arnold, 1993 : 182). Meanwhile, in 1994, along with the reorganisation of the DTI, the economists and statisticians working on science, technology and innovation and certain high technology sectors were brought together with the Assessment Unit to form the Technology, Economics, Statistics and Evaluation (TESE) Directorate. This brought together the two groups of professionals who had been responsible for the evaluation of S&T programmes of the DTI in the past (Barber, 1999a : 35-36).

Of broader significance is the fact that, in 1987, the government created the Science and Technology Assessment Office (AO) within the Cabinet Office. The main function of this body was to encourage and assist all relevant departments to set up assessment system. It pursued this remit by examining each department's existing approach, reporting on it and suggesting improvements. It also published the *R&D Assessment : A Guide for Customers and Managers of Research and Development* in 1989. This has been prepared primarily to help those who are responsible for commissioning or funding R&D for government purposes (Cabinet Office, 1989). This also marked a shift from trying to

count inputs and outputs to trying to manage programmes better. However, it implied that evaluation was merely a part of the 'management information system' needed to implement policy well, and did not address the opportunity to use evaluation strategically as part of the policy formulation process. Nonetheless, the AO and its guidelines gave a major impetus to the use of evaluation across the rest of the government system. In particular, evaluation began to be used as a way to ask questions about science policy as well as industry policy (Guy and Arnold, 1993 : 181). After about two years, the AO was disbanded when the government judged that the practices it promoted had taken sufficient root in departments. In fact, the latter responded to these initiatives with varying degrees of enthusiasm and the demise of the AO was seen in some quarters as a concession to the less enthusiastic. Actually it was probably more that the government felt that departments were too diverse for a uniform approach to be appropriate (Hills and Dale, 1995 : 38).

Nevertheless, the OST has now a somewhat similar preoccupation. It is pressing departments to provide it with performance indicators, sometimes called process measures. It has developed 15 of these, under five headings, which cover many aspects of assessment including evaluation. Examples are the percentage of programmes covered by integrated management systems known as ROAME statements, whether steps to measure wealth creation and quality of life are present, and the extent to which external advisers are brought into the assessment process. In this way, the OST aims to identify the differences between departments and, where possible and appropriate, to achieve standardisation across them. It hopes also to inculcate a stronger ethos of assessment and evaluation generally within departments and research councils, so that eventually it becomes even more a part of their culture. Until this occurs the OST sees its co-ordinating role as crucial (Hills and Dale, 1995 : 38).

More recently, as already noted, the OST has issued a series of guidelines for using scientific advice in policy making process. Of these, the *Guidelines 2000 : Scientific Advice and Policy Making* was a final edition. The guidelines set out some key principles applying to the use and presentation of scientific advice in policy making. They cover the process of identifying issues requiring advice, obtaining the best possible advice from a wide variety of sources, the handling of advice by departments and implementation and review (OST, 2000b : 1-2). As a result, these guidelines are expected to contribute to improving the quality of policy making, through using a lot of evidence-based

information including evaluation findings.

#### 4.2.2.3. Conclusion

It is apparent that the national R&D evaluation system in the UK has been developed in close connection with the development of policy evaluation in general at the national level since the early 1980s. At the moment, in the UK, the evaluation of R&D activities has been recognised as one of the most active and advanced evaluation practices in the public sector, especially in the light of the methodological development and utilisation of evaluation findings. From the late 1990s, the two systems have also influenced each other through, in particular, guidelines such as the *White Paper : Modernising Government* of the Cabinet Office and the OST's *Guidelines 2000 : Scientific Advice and Policy Making*. In sum, the major events and guidance which have had a important role in the process of the developments in the area of policy evaluation and R&D evaluation since the 1980s are summarised in <Table 4-3>.

<Table 4-3> Development of Policy Evaluation and R&D Evaluation in the UK

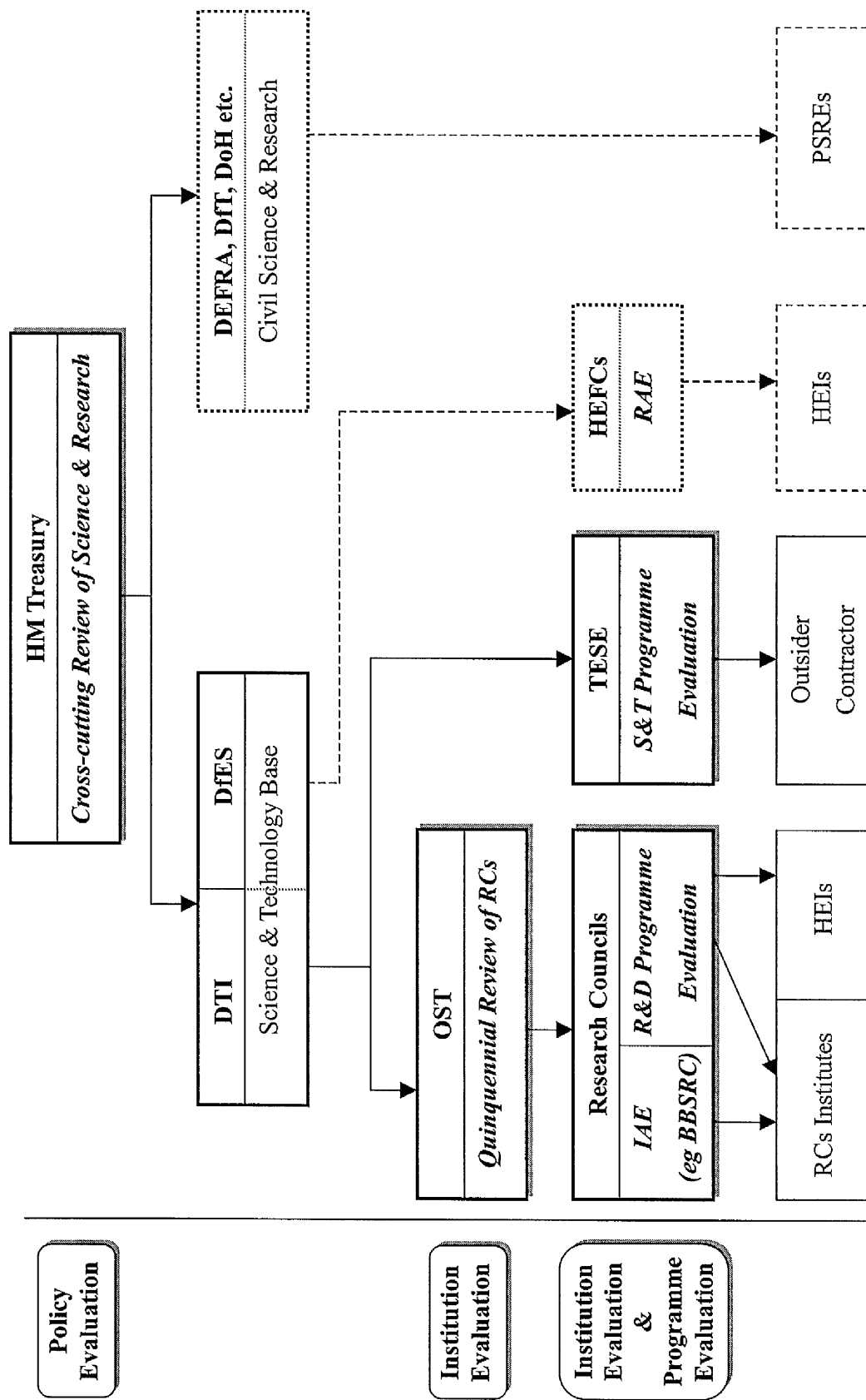
Period Type / Organ.		1980s	1990s	2000s
Policy Evaluation	HM Treasury	- FMI ('82) - <i>Policy Evaluation</i> ('88)	- <i>Economic Appraisal in Central Government</i> ('91) - <i>Green Book</i> ('97)	- <i>Green Book</i> revised ('03)
	Cabinet Office	-	- <i>Modernising Government</i> ('99)	- <i>Guidance on Agency and NDPB Quinquennial Review</i> ('00)
R&D Evaluation	Cabinet Office	- Creation of AO ('87) - <i>R&amp;D Assessment</i> ('89)	-	-
	OST	-	-	- <i>Guidelines 2000</i> ('00)
	DTI	- Creation of STAU (early '80s) - Evaluation of Alvey Programme ('83-'91) - Use of <i>ROAME</i> (late '80s)	- Formation of TESE ('94)	-

\* *Italics* represent the guidelines and related documents in this sector

### ***4.2.3. Current Hierarchy of National R&D Evaluation Systems***

Even though there are many organisations which have carried out evaluation activities in the science and research sector, HM Treasury, the OST, the DTI/TESE and the research councils are the main actors, as noted in Chapter 1. However, each organisation has a different mission in performing the evaluation activities. As a part of spending review every two years, HM Treasury has conducted “cross-cutting reviews of science and research” in 2000 and 2002 (*policy evaluation*). The OST has evaluated the performance of all research councils as a part of a “quinquennial review” (*institution evaluation*). The DTI/TESE has undertaken evaluation of the departmental science and technology (S&T) programmes on the basis of ROAME statements (*programme evaluation*). On the other hand, the research councils have two different missions in their evaluation activities. One is to evaluate all of R&D programmes and their qualities and, sometimes, projects funded by them (*programme/project evaluation*). The other is to evaluate past performance and future plans of research institutes owned or supported by them (*institution evaluation*). Of these, the Institute Assessment Exercise (IAE) conducted by the BBSRC every four years has the reputation of being the most systematic and comprehensive case of the evaluation of research institutes.

Meanwhile, these organisations, which are conducting a different type of evaluation, comprise the upper and lower parts in the national R&D evaluation system as a whole. Thus, the diverse evaluation activities such as institution evaluation by the OST, programme evaluation by the TESE and project evaluation by the research councils should be consistent with the terms and references which result from the “cross-cutting review of science and research”, a type of policy evaluation devised by HM Treasury. Also, both the project evaluation and the institution evaluation conducted by the research councils should be based on the terms of reference under which the OST has been operating in the area of science and technology policy. In fact, the current hierarchy of the national R&D evaluation system in the UK can be seen in (Figure 4-3).



(Figure 4-3) Hierarchy of National R&D Evaluation System in the UK



#### ***4.2.3.1. Evaluation System for Science and Research Policy***

HM Treasury has the general task of monitoring the effectiveness of all government programmes and initiatives. These evaluations will inform future policy development and spending priorities. HM Treasury has firstly conducted a Comprehensive Spending Review (CSR) in 1998 and it also has done Spending Reviews (SRs) in 2000 and 2002 respectively. While the CSR was a zero-based approach to spending, the SRs are based on existing spend and assess changes from existing level.

Spending Reviews set out the key improvements in public services that the public can expect over a three-year period. They include a thorough review of departmental aims and objectives to find the best way of delivering the government's objectives. Key features of the Spending Reviews are (1) department's taking a more fundamental look at the effectiveness of spending, rather than falling into the familiar annual exercise of bidding for extra resources, (2) a focus on objectives and outcomes, allowing the government to look across the work of departments to find cross-departmental solutions to achieve its objectives and to encourage joint working, (3) directing spending towards the government's key objectives, education, health, transport and law and order, in new three-year spending plans and (4) the introduction of Public Service Agreements (PSAs),<sup>2</sup> detailing the exact outcomes departments will deliver with the money provided, such as better health and higher educational standards. HM Treasury monitors progress against PSA targets and reports in detail in annual departmental reports, giving Parliament and the public the opportunity to evaluate progress (HM Treasury, 2000b).

Consequently, HM Treasury has a wide interest in ensuring that the UK science and technology sector has an economic environment in which it can thrive and in which innovation is positively encouraged. Furthermore, it ensures that value for money is

---

<sup>2</sup> PSAs are commitments to the public on what they can expect for their money and each agreement sets out explicitly which minister is accountable for delivery of targets underpinning those commitments. The publication of PSAs is the beginning of the process. Every department is expected to ensure that the targets are delivered. The next step after establishment of PSA target is the announcement of a Service Delivery Agreements (SDAs) which set out how the targets will be achieved. In particular, the targets for the science and research expenditures are set for the DTI, as the department which is ultimately responsible for the research councils and SET base. The Secretary of State for Trade and Industry is accountable to Parliament for the achievement of the objectives set out in the PSA and SDA, for the delivery of the targets and for the resources consumed in the process (HM Treasury, 2000b). For more detailed PSAs and SDAs of the DTI, see "Chapter 11 : Department for Trade and Industry" in the *2000 Spending Review Public Service Agreements White Paper (2001-2004)* (HM Treasury, 2000c).

maximised in spending on science and technology by government departments. So, it monitors departmental spending on science and innovation through spending teams, which oversee the work of each department. In most cases, there is a single spending team for all large departments. Some spending teams oversee the work of two or more smaller departments. The “Enterprise Team” acts as the spending team for the DTI, which is responsible for science, technology and innovation at the national level, and has the lead responsibility within HM Treasury for science, innovation and enterprise. The Enterprise Team sits within the Enterprise and Growth Unit in the Finance, Regulation and Industry Directorate. It aims to assist in the creation of a more enterprising society, in which more firms of all kinds, both new and existing, have high ambitious and can achieve their potential. It does this by, for example : (1) ensuring that public funding concentrates on excellence and rewards exploitation, (2) encouraging universities and public sector research establishments (PSREs) to increase their ability to manage and exploit intellectual property effectively and (3) encouraging more innovation in business (HM Treasury, 2001a : 3).

As a part of the spending review 2000, a “crossing-cutting review of science and research” was carried out, chaired by a senior HM Treasury official with representatives from other government departments and external advisers from the universities and private sector. The review examined publicly funded science and engineering research and made a number of recommendations, which were taken forward in the Spending Review 2000 and the Science and Innovation White Paper 2000, *Excellence and Opportunity : A Science and Innovation Policy for the 21st Century*, produced by the DTI in July 2000 (HM Treasury, 2000a).

A further cross-cutting review of science and research, one of the seven initial cross-cutting reviews that will contribute to the Spending Review 2002, was announced in June 2001. The science and research cross-cutting review includes a review of (1) funding of the UK science base, and (2) the effectiveness of departments’ own science and research programmes to ensure that they deliver maximum long term benefits to the economy and quality of life. The full terms of references are to consider how to maximise the benefits provided by public spending on science and research to the UK economy and quality of life. In particular, the review of “science funding by the OST/DTI and the DfES” (1) reviews current funding mechanisms and levels, and identifies the priorities for

resources across the funding streams held by the OST and the DfES, (2) reviews current funding mechanisms for knowledge and technology transfer from the science and engineering base and mechanisms in order to promote their adoption by business, and identifies priorities for the future use of resources, (3) takes stock of studies commissioned or reported since the last spending review, with their potential implications for resources, including 'The Transparency Review on the Costing of University Activities', 'The Robert Review into the Supply of Scientific Skills for Business R&D', 'The Review of Investment in University Research Infrastructure', 'The Quinquennial Review of the Six Grant Awarding Research Councils', 'The Review of Business Support Schemes in the DTI' and 'The HEFCE/OST Higher Education Business Interaction Survey'. Another aspect of work, the review of "civil research commissioned by government departments" (1) assesses the scientific and technological capabilities of government departments with significant research needs, reviews current and planned levels of activity and compares both capabilities and activity levels with likely scenarios over the next decade, and identifies priorities, (2) identifies proposals for improving the effectiveness and value for money of civil research commissioned by departments in order to maximise the impact of available resources, and increases the contribution made to the government's broad goals for the economy and quality of life, including mechanisms for knowledge and technology. For this review, the Minister for Science and Innovation in the DTI is the ministerial lead, and has been supported by a team of officials from HM Treasury, the OST and the DfES (HM Treasury, 2001b).

#### ***4.2.3.2. Evaluation System for Research Institutions***

Regarding the evaluation of institutions in the science and technology sector in the UK, there are two different types in terms of evaluation purpose, evaluation object and so on. One is a "quinquennial review" in which the OST reviews the rationality and appropriateness of the research council's mission, role and so on, and evaluates its past performance every five years. The other is a "regular evaluation on the research councils' institute" in which some research councils, for example the BBSRC, the MRC, the NERC and the PPARC, evaluate the research performance of research institutes or research centres owned or supported by them and their research quality. As already noted, since the BBSRC's four yearly Institute Assessment Exercise (IAE) has been conceived as a best example among the institution evaluations by the research councils, so I shall choose

this IAE for a case study.

### **(1) Quinquennial Review of the Grant Awarding Research Councils by the OST**

In the UK, agencies and Non-Departmental Public Bodies (NDPBs) have been at the heart of the programme to modernise government. They provide key services which are central to the activities of government and affect the lives of all citizens. The government needs to review organisation and performance regularly to make sure it has delivered what citizens need. So, quinquennial reviews of NDPBs are considered a key component of government's commitment to improving the quality and effectiveness of public services. These reviews are required to address the alternative options for the provision of the service, the way that the service fits into wider governmental objectives and the extent of any synergies with other service providers (Cabinet Office, 2000a : 2 ; DTI/OST, 2001b : 1).

The guidance for the conduct of reviews (*How to Review Agencies and Non-Departmental Public Bodies to Improve the Quality and effectiveness of Public Services*) sets out a two stage format in which stage one is about 'getting the organisation right', while stage two addresses 'improving performance'. Each stage is required to be completed within three months. The guidance sets out a number of issues that should be considered in stage one :

- (1) How the agencies and NDPBs contribute to wider departmental and government objectives.
- (2) What links the agencies and NDPBs have, or should develop, with other bodies.
- (3) Past performance and responsiveness to customer requirements
- (4) Evidence of good practice, such as EFQM Excellence Model and IiP
- (5) Consideration of alternative delivery mechanisms (options) covering : abolition, continued NDPB status, market testing, merger or rationalisation, privatisation and strategic contracting out (Cabinet Office, 2000a).

Both the DTI and the OST, who are responsible for this review, initiated a "Quinquennial Review of the Six Grant Awarding Research Councils", including the BBSRC, the EPSRC, the ESRC, and MRC, the NERC and the PPARC in February 2001. The intention of this review was to consider the Councils both individually and as a group. The White Paper on Science and Innovation (*Excellence and Opportunity* ; Cmd 4814) and the

Knowledge Economy (*Opportunity for All in a World of Change* ; Cmd 5250) provide the context for it. Its terms of reference for this review set the following more detailed objectives for a quinquennial review :

- (1) The first stage examines the role and organisation of the Councils by reference to their Charters and missions, the evidence of work already undertaken to review and improve performance, current best practice for NDPBs and the Councils' contribution to the delivery of government policy for publicly funded scientific research and training in the context of recent White Papers. It also sets the detailed terms of reference for the second stage.
- (2) The second stage examines the efficiency and effectiveness of the Councils' operation. It also looks for further opportunities for improving performance by reference to structures, aims and objectives, the roles of funded institutes, performance targets and service standards, the use of process and new technology, delegated authority and accountability and changes to the operating environment since the Councils were awarded their Royal Charters in 1994, including the *Modernising Government* agenda (DTI/OST, 2001b : 1-2).

## **(2) Institute Assessment Exercise of the BBSRC**

The BBSRC supports research and training in over 60 universities and in eight sponsored research institutes. Some 35% of BBSRC funds are allocated to its institutes, which deliver fundamental and strategic research in the non-medical life sciences. The eight BBSRC-sponsored institutes have distinctive missions and have played a significant role in meeting the objectives of science.<sup>3</sup> The BBSRC provides funding for its eight institutes through Core Strategic Grants (CSG) provided on a four-year basis following extensive review of past performance and future plans. This is the "Institute Assessment Exercise" (IAE) (BBSRC, 1999a : 3, 26).

Following the first IAE carried out in 1996-1997, in April 1999, the BBSRC Council agreed that a further and modified exercise would be conducted in 2001. The outcome of this exercise informed the Councils' allocation of the CSG to the BBSRC-sponsored

---

<sup>3</sup> For more detailed, key features of the institutes are : (1) distinctive and focussed research missions, (2) strategic and multidisciplinary research orientation with a longer term perspective, (3) user community input, in particular at governing bodies and advisory groups, (4) knowledge and technology transfer, including intellectual property, know-how and public good information and (5) special facilities, for example disease containment, animal houses and glasshouses (BBSRC, 2002a).

institutes and Horticulture Research International for the four financial years from 2002-2003 to 2005-2006. The 2001 IAE comprised the following four components : (1) assessments by expert panels of the quality of science (QS), (2) an assessment by a panel of knowledge transfer (KT), (3) an assessment by the BBSRC Office of competitive funding (CF) and (4) an assessment of the institute as a whole by a visiting group (VG), which included an assessment of its strategic relevance and the advanced training provided (BBSRC, 1999a : 1).

#### **4.2.3.3. Evaluation System for S&T and R&D Programme**

In the evaluation of the S&T and R&D programmes, the DTI and the research councils are the main actors responsible for the “Innovation Budget” and the “Science Budget”, respectively. The DTI has a well-established system of evaluation, which starts with the genesis of a new initiative and runs until some years after the initiative has ended. The foundation of this system is the ROAME<sup>4</sup> statement for expenditure programmes. All programmes, whatever the level of funding, must have a ROAME statement or its equivalent in order to record the justification for the expenditure. Small programmes generally require a less comprehensive ROAME statement than those which seek significant sums of public funding.

Meanwhile, the main responsibility for evaluating the Science Budget funded research has been delegated to the research councils. They conduct *ex ante* appraisal of research proposals, a minority of which are funded, and then *ex post* evaluation of the outcome of research, through examination of the end of programme report. At project and programme level, the peer review system grades the scientific quality of research by the evaluation of final research reports (DTI, 2001a : 32).

#### **(1) S&T Programme Evaluation by the DTI**

The DTI is the main sponsoring department for UK industry and services. However, some

---

<sup>4</sup> The ROAME is the acronym of Rationale, Objectives, Appraisal, Monitoring and Evaluation. The function of the ROAME is to set out the reasons (1) why government should act at all (Rationale), (2) what the objectives for the programme/activity are to be (Objectives), (3) how individual projects/activities will be selected for inclusion (Appraisal), (4) how the programme and its constituent activities will be monitored (Monitoring) and (5) the arrangements, together with some suggestions as to the focus of an evaluation of the programme (Evaluation). The ROAME therefore provides a framework for future evaluation.

important industries have been sponsored by other departments. For example, the DoH sponsors the pharmaceutical industry, the DEFRA sponsors the food and environmental industries, the DfT supports the transport industries, and the UK defence businesses are mainly in contact with the MoD which is their principal customer, though the DTI takes an interest in them as well. Support for all other industries and services are the responsibility of the DTI (Barber, 1999a : 1).

Since the DTI was an early leader amongst the UK government departments in carrying out systematic *ex ante* appraisal and *ex post* evaluation of S&T programmes and is still seen as an example of best practice, its approaches have spread to other departments. Thus, to a considerable extent, the DTI's method and process of evaluating S&T programmes can be seen as typical of the UK approach to evaluating government intervention to support industry and commerce (Barber, 1999b : 33).

The DTI provides support for the science, technology and innovation activities of the sectors which it sponsors from four different budgets. These are (1) the Innovation and Technology Support Budget, (2) the Aerospace Budget, (3) the Space Budget and (4) the Energy Budget, covering the various energy sectors. The Technology, Economics, Statistics and Evaluation (TESE) Directorate is responsible for the evaluation of programmes and support under the first three of these budgets. The evaluation of S&T components of the Energy Budget is undertaken elsewhere in the DTI (Barber, 1999b : 1-2).

In the DTI, the evaluation of S&T programmes which is one of the major elements in the ROAME statement is undertaken to : (1) demonstrate that a programme has met its objectives and satisfied its rationale so that tax payers' money has been well spent, (2) improve the effectiveness of current programmes, (3) inform the design of future programmes, (4) increase the knowledge base on which the S&T policies are founded (Barber, 1999a : 12). The actual S&T programme evaluations have been carried out by the Technology and Standards Assessment Unit (AU) which is a part of the TESE. In carrying out evaluations, the AU draws on the expertise of the professional economists from other parts of the TESE. Also, advice on surveys and statistical matters is available from the professional statisticians (Barber, 1999a : 6).

In addition to the DTI, other civil government departments such as the DEFRA, the DfT and the DoH have run their own R&D evaluation systems supporting science and technology. As already described, their evaluation systems have been greatly influenced by the DTI's experience and development of the theories. The R&D evaluation systems of the civil departments are summarised in <Table 4-4> on the basis of the information acquired from the existing literature (Barber, 1999a ; Barber, 1999b ; Barber, 2001 ; Hills, 2001).

**<Table 4-4> R&D Evaluation Systems in the Civil Departments in the UK**

<b>Depart. Function</b>	<b>DTI (excluding OST)</b>	<b>DfT (former DETR)</b>	<b>DEFRA (former MAFF)</b>	<b>DoH</b>
<b>Major R&amp;D Performer</b>	-outside contractor	-outside contractor	-largely carried out by its agencies	-N.A.
<b>ROAME Statement</b>	-strongly applied	-variably and partly applied -not use at the appraisal	-centrally applied	-not use
<b>Responsibility (especially in <i>ex post</i>)</b>	-committee (senior finance officer, policy makers, evaluators)	-the Departmental Research Strategy -fully consistent with OST guidance	-Chief Scientist's Group (CSG)	<b>*not institute formal <i>ex post</i> evaluation</b>  *concentrate on <i>ex ante</i> appraisal  *encourage publication of research results in academic journals  *research process reviewed by advisory committees
<b>Purpose</b>	-research excellence -policy contribution	-research excellence	-quality of research -value to its policies -directing future research	
<b>Type</b>	-output evaluation -impact evaluation	-output evaluation -impact evaluation	-N.A.	
<b>Timing</b>	-just before completion	-on completion -in long periods, at least every 5 years	-soon after completion -in longer case, every three years	
<b>Level</b>	-policy -programme	-policy -programme	-programme	
<b>Evaluator</b>	-internal evaluators -a few external evaluators	-both internal and external evaluators	-external researchers	
<b>Unit</b>	-separate unit (TESE)	-three 'in-house consultancy units'	-none	
<b>Method</b>	-interview (face to face, telephone, questionnaire etc.)	-similar to DTI -assessment of achievement of objectives	-similar to the DTI's and the DfT's	
<b>Utilisation</b>	-progra. management -programme/policy design -budget allocation -value for money	-distribute executive summary to heads concerned	-N.A.	

Sources : Barber (1999a), Barber (1999b), Barber (2001), Hills (2001).



## **(2) R&D Programme Evaluation by the BBSRC**

The BBSRC evaluates all its major research programmes, in order to assess (1) whether they have achieved their aims and objectives, (2) whether there have been any major contributions made to the subject area as a results of the programme, (3) whether there have been any benefits gained from the programme being co-ordinated centrally by the BBSRC, (4) whether there has been any added value in awarding grants through an initiative programme rather than through normal responsive mode grants and (5) the level of the outputs from the programme. This information has been used to identify priority areas for future funding.

The BBSRC issues a questionnaire to all grantholders, asking for their views on issues such as whether the programme has met its objectives, how well it was managed by the BBSRC, details of new contacts/collaborations, further funding applied for or received etc. If the programme had workshops associated with it, then the BBSRC asks for grantholders' views on their usefulness for networking. Grantholders are also asked to give details of their research outputs from the programme specifically funded by the grant, including refereed publications, major conference presentations or abstracts and activities to promote public understanding of science. The information gained from questionnaires is then amalgamated with details of outputs from final report forms, for example publications, gene sequences, etc., in a report and bibliometric analysis.

A review panel is organised by the research committee which coordinates the programme in the BBSRC's Science and Technology Group. The panel consists of several members of the original steering group plus other independent members, with the BBSRC staff as observers. As well as considering the questionnaire responses and bibliometric analysis, the panel also reviews the scores of individual grants from final reports. For some programmes, progress reports given part of the way through the programme are available and also considered. The conclusions of the review are submitted to the next meeting of the relevant research committee to consider the way forward (BBSRC, 2001b).

### ***4.2.4. Guidance and Consultation Documents for R&D Evaluation***

UK central government departments such as HM Treasury, the Cabinet Office and the OST have issued some important guidance and published consultation documents on the

execution of both policy evaluation and R&D evaluation. While some of them have been replaced by revised or new versions, some are still operative in this field. Along with the central government, each executive department, including the DTI and the BBSRC, also has its own guidelines, which are consistent with the terms of reference of the central governments guidelines. These guidance and consultation documents are summarised in <Table 4-5>. In this part, I shall give a brief explanation of the current guidelines and consultation documents at the level of both central governments and executive departments.

**<Table 4-5> Guidance and Consultation Documents for R&D Evaluation**

<b>Type of Evaluation</b>		<b>Guidance &amp; Consultation Document</b>
<b>Evaluation of S&amp;T and R&amp;D Policy</b>		<ul style="list-style-type: none"> <li>- Policy Evaluation (HM Treasury, 1988)</li> <li>- R&amp;D Assessment (Cabinet Office, 1989)</li> <li>- Economic Appraisal in Central Government (HM Treasury, 1991)</li> <li>- The Green Book (HM Treasury, 1997)</li> <li>- Modernising Government (Cabinet Office, 1999)</li> <li>- Guidelines 2000 (OST, 2000)</li> </ul>
<b>Evaluation of Research Institution</b>	<b>Evaluation of Research councils</b>	<ul style="list-style-type: none"> <li>- R&amp;D Assessment (Cabinet Office, 1989)</li> <li>- Modernising Government (Cabinet office, 1999)</li> <li>- Guidance on Agency and NDPB Quinquennial Review (Cabinet Office, 2000)</li> </ul>
	<b>Evaluation of Research Institutes (e.g. by BBSRC)</b>	<ul style="list-style-type: none"> <li>- R&amp;D Assessment (Cabinet Office, 1989)</li> <li>- IAE 2001 Guideline (BBSRC, 1999)</li> </ul>
<b>Evaluation of S&amp;T / R&amp;D Programme</b>	<b>Evaluation of S&amp;T Programme (e.g. by DTI)</b>	<ul style="list-style-type: none"> <li>- Guidance on Preparing Evaluation Plans (DTI, 2000)</li> </ul>
	<b>Evaluation of R&amp;D Programmes (e.g. by BBSRC)</b>	<ul style="list-style-type: none"> <li>- A lot of internal documents</li> </ul>

#### ***4.2.4.1. R&D Assessment : A Guide for Customers and Managers of Research and Development (Cabinet Office, 1989)***

This guide has been written primarily to help people who are responsible for

commissioning or funding R&D for government purposes. It describes (1) the different steps and processes of R&D assessment, (2) when they are needed, (3) how they should be tied to policy making, (4) the importance of involving the right people and (5) the use of particular techniques and indicators. It also suggests that most funding bodies need to take a further step by having their own assessment guidelines that take account of their own and the departmental objectives, how R&D fits in and the types of R&D they fund.

The guide is in two sections. Section One sets out good practice as it applies to most types of R&D, and serves as a general introduction. In principle, it groups government funded R&D activities into three main categories : (1) basic and strategic science, (2) support for development of potentially exploitable technology and (3) R&D which government supports to fulfil its policy making and statutory duties. It indicates that R&D assessment normally comprises three different types of activities : appraisal (before R&D starts), monitoring and review (during R&D) and evaluation (after R&D). Following the classification of R&D and the definition of R&D assessment, it explains the basic factors and principles which cover appraisal, monitoring and review, and evaluation, including the actors, the assessment process, assessment techniques, and assessment cost. Along with the assessment of programmes and projects, it also suggests that we need to think about the institutions which do the work and how to assess their effectiveness.

Section Two provides more details of the assessment for different types of R&D. For each of the three types of R&D, it suggests a process, criteria and method of assessment which are in principle similar but to a certain extent different. In particular, it recommends that evaluation should consider three types of evaluation criteria : effectiveness, efficiency and follow-up action. For example, in the area of “basic and strategic science”, even though the scientific quality of proposals, the utility and the possible social and economic benefits are used as the general assessment criteria, the main thrust appears to be towards scientific excellence. In the case of “support for policy making and statutory duties”, the output may take the form of technical advice, a contribution to a code of practices and the development of tools to help with enforcement or the setting of standards. Finally, “support for technology and procurement” is mission-oriented R&D. Since its objectives are more clear and specific, the assessment can be objective and place relatively little reliance on subjective judgement.

#### **4.2.4.2. *The Green Book : Appraisal and Evaluation in Central Government (Treasury Guidance) (HM Treasury, 1997)***

The principles of central government appraisal and evaluation in this guide apply to any activity, whether it be a project, programme, or policy, which entails spending or saving money or otherwise changing the ways in which resources are used. Guidance is provided on basic principles and on many broader technical issues for which there is a common interpretation of good practice.

This guidance comprises four chapters and some annexes. Chapter One is a general introduction. Chapter Two and Three provide a background to appraisal and evaluation for those who may be new to, or who have only limited knowledge or experience in this field. Chapter Four provides advice on technical matters and is aimed more at those who undertake appraisal and evaluation. The document is not intended for reading from cover to cover in one sitting. The chapters, annexes and appendices provide progressively more detail that the reader can choose for his or her particular needs.

This edition supersedes HM Treasury's first edition, *Economic Appraisal in Central Government*, issued in 1991 and has been extensively revised. It provides more material on evaluation. It gives greater emphasis to the appraisal and evaluation of environmental impacts. It takes account of developments in the treatment of these impacts and other costs and benefits which are not easy to value. It takes account of developments in the use of private finance and, related to this, provides a more thorough coverage of the treatment of risk and uncertainty. The section on industrial and regional programmes has been extended to cover a broader range of programmes aimed at raising economic activity. The guide confirms the use of a six-percent real discount rate in most circumstances. The technical annex contains a more extensive discussion of the derivation of the discount rate in terms of the cost of capital and time preference.

This guide is not a substitute for the technical guidance developed by departments. For many areas, specific departmental guidance is also required, consistent with this book, and geared directly to departmental needs. Previous editions have been used in some parts of local governments and other outside central governments. Such use is welcome, although the guidance has been prepared with the needs of central government in mind.

This guide also suggests that training in appraisal and evaluation is important and this guide can provide a positive contribution to this.

#### **4.2.4.3. *Modernising Government (Cabinet Office, 1999)***

The White Paper, *Modernising Government*, is central to the government's programme of renewal and reform. In line with the government's overall programme of modernisation, *Modernising Government* is modernisation for a purpose : to make life better for people and businesses. This white paper is a long-term programme of improvement. In order to ensure that government is both inclusive and integrated in policy making and its implementation, the white paper has three aims : (1) ensuring that policy making is more joined up and strategic, (2) making sure that public service users, not providers, are the focus, by matching services more closely to people's live and (3) delivering public services that are efficient and of high quality.

The government is centring this programme on the following five key commitments :

- (1) Policy Making : The government will be forward looking in developing policies to deliver outcomes that matter, not simply reacting to short-term pressure. For this, it will identify and spread best practice through the new Centre for Management and Policy Studies, bring in joint training of ministers and civil servants and introduce peer review of departments.
- (2) Responsive Public Services : The government will deliver public services to meet the needs of citizens, not the convenience of service providers. So, it will deliver a big push on obstacles to joined-up working, through local partnerships, one-stop shops, and other means, and will involve and meet the needs of all the different groups in society.
- (3) Quality Public Services : The government will deliver efficient, high quality public services and will not tolerate mediocrity. It will review central and local government department services and activities over the next five years to identify the best supplier in each case, set new targets for all public bodies, focusing on real improvements in the quality and the effectiveness of public services, and monitor performance closely so that government strikes the right balance between intervening where services are failing and giving successful organisations the freedom to manage.
- (4) Information Age Government : The government will use new technology to meet the

needs of citizens and business, and not trail behind technological development. In other words, it will develop an information technology (IT) strategy for government, which will establish cross-government co-ordination machinery and frameworks on such issues as the use of digital signatures and smart cards, websites and call centres and benchmark progress against targets for electronic services.

- (5) Public Service : The government will value public service, not denigrate it. For this purpose, it will modernise the civil service, revise performance management arrangements, tackle the under-representation of women, ethnic minorities and people with disabilities, build the capability for innovation and establish a public sector employment forum to bring together and develop key players across the public sector.

Regarding R&D evaluation activity, of the five commitments for modernising government, “quality public services” are most relevant to the evaluation of the S&T and R&D sectors, in particular, the evaluation of research councils, one type of NDPBs, which are main suppliers of science, engineering and technology in the public sector. Under the terms of this white paper, *Guidance on Agency and NDPB Quinquennial Review* was published in 2000, in order to give detailed guideline to the departments that are responsible for the management and evaluation of the agencies or NDPBs at regular periods.

#### **4.2.4.4. *Guidelines 2000 : Scientific Advice and Policy Making (OST, 2000)***

*Guidelines 2000* sets out some key principles applying to the use and presentation of scientific advice on policy making. They should be read in conjunction with guidance on related aspects of policy making, such as the *Modernising Government* programme, the Health and Safety Executive (HSE) publications on risk assessment, risk management and risk communication, and other publications.

These guidelines apply to advice and research in science, engineering and technology, although aspects may usefully be applied to a broader range of issues involving other disciplines. They cover the processes of (1) identifying issues requiring scientific advice, (2) obtaining the best possible advice from a wide variety of sources, (3) handling of advice by departments and (4) implementation and review.

The guidelines are intended primarily for the use of UK government departments, although they will be relevant to any body that uses scientific advice in informing policy and to all scientific disciplines falling within the definition of SET used by the Office of National Statistics for its Government R&D Survey. These also apply to all areas in which scientific advice is required and whatever the sources of scientific advice to government : whether it is from in-house or sole external experts, standing or ad-hoc advisory committees, contract research from academia, industry or commerce and independent research or elsewhere. However, it is particularly important that they are followed carefully where the issues are sensitive, for example where there is significant scientific uncertainty, a range of scientific opinion, or implications for public policy. Departments should use their judgement to apply these guidelines in a manner which is proportionate to the nature and scale of the issues involved.

In broader terms, these guidelines are more concentrated on the process and quality of policy making. However, regarding R&D evaluation, *Guidelines 2000* suggests that in order to improve the quality and rationality of its policy making, each government department should use as much evidence-based information, for example evaluation findings as possible. Meanwhile, in December 2001, to guarantee the implementation of these guidelines at the level of executive department, the OST published the report, *Scientific Advice and Policy Making : Implementation of Guidelines 2000*, which dealt with the implementation process and results in 19 government departments and agencies (OST, 2001a).

#### ***4.2.4.5. Guidance on Agency and NDPB Quinquennial Review : How to Review Agencies and Non-Departmental Public Bodies to Improve Quality and Effectiveness of Public Services (Cabinet Office, 2000)***

Agencies and NDPBs are at the heart of the long-term programme to modernise government. They provide key services which are central to the activities of government and affect the lives of all citizens. So, this guidance sets out a radically new approach to the way in which government departments should review agencies and executive NDPBs so that they deliver better services to the public that are easily accessible and simple to use. Agency and NDPB reviews are a major component of a programme of continuous

improvement based on the five principles : challenge, compare, consult, compete and collaborate.

The government is using these principles to build on its Best Value approach to local government, and complement the existing scrutiny of central government carried out by the National Audit Office and by Parliament. These principles are also being acted upon by setting clear targets through departmental and cross-cutting Public Service Agreements and through the Better Quality Services initiative, reviewing all central government activities and functions to identify the best supplier.

This guidance comprises three parts. Part One is designed to give ministers and departments an overview of what to look for in reviewing executive agencies and NDPBs. Part Two provides guidance on general issues that will be relevant to most reviews. Finally, Part Three provides more detailed guidance on specific issues and aspects to look for in the forms of fact sheets. Meanwhile, this guide has three main aims : (1) to ensure that all agencies and NDPBs are subject to regular and challenging review, taking into full account the views and needs of their customers, staff and their trade unions, and other stakeholders, (2) to ensure that government puts in place the right structures for these bodies, so that they deliver its agenda effectively, and work in an effective and joined-up way, (3) to provide a strong focus on improving their future performance, including the scope for more partnership working, and better use of new technology.

The key features of the new approaches adopted in this guidance are as follows :

- (1) Regular Review : Reviews of agencies and NDPBs should, as now, take place at least every five years. Where it makes sense, review may be brought forward, for instance to fit in with the timing of the public spending review, or be delayed, for example, to allow the results of other work to inform the review. In such cases, prior approval of the Cabinet Office and HM Treasury ministers is required for agencies and large NDPBs with a staff of 100 or more. In addition, it is open to agencies and NDPBs to fulfil their commitment to the Better Quality Services review through their five yearly review. In most cases, where the agency and NDPB is not too large, quinquennial reviews can be structured to achieve this.
- (2) The Bigger Picture : Just as agencies and NDPBs do not exist in isolation from the rest of government, so they should contribute to the delivery of wider departmental and



governmental objectives. In thinking about how an Agency or NDPB fits into what the government wants to achieve more widely, the starting point is likely to be the departmental and cross-cutting Public Service Agreements (PSAs) which set out the aims and objectives of each department along with the targets identified in order to achieve them.

- (3) All Options Should Be Considered : Agency reviews should consider all options including : abolition, continued agency or NDPB status, market testing, mergers or rationalisation, privatisation and strategic contracting out. Each of the organisational options should be assessed equally on its merits. The arrangements for NDPBs are different. NDPB reviews will continue to be required to consider first whether the function is required at all, and if it is, whether an NDPB is the best option for its delivery. This reflects the government's view that it should seek to reduce the number of NDPBs. In practice, it may sometimes be the case that the best means of delivering services and functions in the future is a combination of more than one of the options mentioned.
- (4) Joined-up Government : The starting point for review will remain the single agency or NDPB. However, the new guidance makes it clear that there may be occasions when it makes more sense to look at a number of organisations at the same time, either as part of the same review (joint) or using two or more reviews running concurrently (parallel). In addition, each review will consider the scope for closer working between different parts of the public sector, for example, central and local government, as well as with the private sector.
- (5) New Focus on Future Improvements : The guidance makes it clear that if ministers agree that continued agency or NDPB status is the right organisational structure, the second phase of the review should be forward-looking, providing a much greater emphasis on how service and functions can be provided more effectively in the future, including the use of new technology.

This guidance also sets out in some detail how reviews are to be carried out. Key points include :

- (1) A Two-stage Approach : In line with the new emphasis on agencies' performance, the guidance provides for a two-stage approach for agencies (NDPB reviews are already undertaken in two stages), under which the organisational options are considered first and then, if agency or NDPB status is confirmed, a forward-looking examination

takes place of how to improve performance. This replaces the previous three-stage approach for agencies. In practice, there is likely to be some iteration between the two stages.

- (2) Shorter Review : The scale of a review should be commensurate with the job to be done. Larger agencies and NDPBs, and those with more complex activities, will naturally require more work than smaller and simpler organisations. In addition, those with an interest must be given sufficient time to contribute. However, long running reviews are counterproductive for the staff working in the agency and NDPB, and delay service improvements. Departments therefore should aim to complete stage one within three months at most (from start to presentation of recommendations to departmental ministers) and stage two within a further three months (from agreement on the conclusions of the first stage).
- (3) User and Staff Focus : Finding out what customers want is a key part of the *Modernising Government* programme and is vital in order to achieve the drive for delivering services which provide what customers want. The customers' view should inform both stages of the reviews, as should the views of agency or NDPB staff, including those at the frontline, and their trade unions. Allied to this, departments should consider including a member of staff from the agency and NDPB on the review team to aid the review's understanding of the organisation's business and context in which it operates. Departments and agencies/NDPBs should agree and plan well in advance what information will be needed for the review, and how it will be collected. Information from other sources, for example, existing surveys can be used where it is relevant and sufficiently current. Publicity should be given about reviews in good time to enable all with an interest to contribute.
- (4) External Challenge : Departments, rather than agencies or NDPBs, are responsible for carrying out the review on behalf of their ministers. This is in the interests of objectivity and as a challenge to internal thinking. Departments must make sure what internal level of resources they have to devote to this task. As part of this, it will normally be the case that a reviewer from outside the department will be involved in at least Stage One of the review and, where appropriate, Stage Two. Having external representatives on a review's Steering Group can also be an effective way of providing independent input.
- (5) Agency/NDPB Involvement : Although not responsible for carrying out the review, the agency or NDPB should be closely involved. The terms of reference for reviews

should be made available to staff and their trade unions, who should be given the opportunity to put their views to those undertaking the review. They should be informed of minister's decisions on the result of the review and consulted on its implementation.

- (6) Good Practice : The regular review process provides an excellent opportunity to help disseminate good practice. There is therefore a requirement on departments to provide the Cabinet Office with a list of key lessons learnt and good practice identified at the end of each review, and to include it in the final review report. This can cover both the conduct of the review and any point relevant to improving the delivery of services.

#### ***4.2.4.6. Guidance on Preparing Evaluation Plans (DTI, 2000)***

As already noted, the DTI has played a major role in the development of both policy evaluation and R&D evaluation using the ROAME statements. As a part of this, it has also published some important guidelines and consultation documents in relation to policy evaluation. Of these, *Guidance on Preparing Evaluation Plans*, issued in March 2000, is one of the current guidelines for the evaluation of diverse policies and programmes which have been funded and sponsored by it.

This guidance has been prepared to assist evaluators and programme managers in the development of plans for the evaluation and monitoring of DTI policies and programmes. It sets out the core elements which should be covered in evaluation-planning papers. One task of the ES Central Evaluation Team is to provide guidance on preparation of an evaluation strategy or planning paper. The planning paper should be agreed with them before proceeding with the evaluation or inviting external consultants to tender.

Generally speaking, evaluation is needed to improve the success of policies in achieving departmental objectives, to assist resource allocation and to provide accountability. As in other departments, in the DTI, the importance of evaluation has been recently enhanced by the introduction of Public Service Agreement (PSA) targets and the *Modernising Government* agenda, with its emphasis on evidence based policy. The scope of evaluation in the DTI encompasses the full range of policy instruments, including regulation and staff based activities, as well as expenditure-based programmes. Monitoring and

Evaluation are covered in the “M” and “E” sections of the Department’s ROAMEF (Rationale, Objectives, Appraisal, Monitoring, Evaluation and Feedback) guidance.

In the DTI, evaluation-planning papers have a key role in ensuring the quality, transparency and policy relevance of the department’s evaluation process. They should set out the detailed research requirement for the evaluation, drawing on the evaluation and monitoring elements of the programmes’ ROAMEF statement. The aim is to present clearly the objectives and expected outputs of the evaluation, the reasons why it is needed, how the evaluation timing fits into the policy process, what fundamental questions it seeks to address, what methodology will be used, and the expected costs. Discussion of the papers in the Evaluation Methodology Group provides an opportunity for peer review among evaluators, and ensures that the proposed approach to evaluation meets the department’s criteria for quality, contents and consistency. It also provides an opportunity for the programme managers to ensure that their information needs and their knowledge of the policy area are reflected appropriately in the research design.

Once agreed, evaluation-planning papers serve as the specification for evaluation research, and with the exception of information on evaluation costs, it should also be attached to the “Invitations to Tender” for external evaluation contractors.

#### ***4.2.4.7. Guidelines to Institutes on the Procedure and Information Requirements for Submissions to the 2001 Institute Assessment Exercise (IAE) (BBSRC, 1999)***

In order to achieve the scientific objectives, the BBSRC sponsors eight strategic research institutes, has set up six structural biology centres, and funds a number of other research centres. Of these, eight research institutes are the fundamental and strategic research performers in the agriculture and food areas. Regarding funding to the research institutes and some other research programmes, the BBSRC has conducted a regular review and evaluation of past performance and future plans, known as the “Institute Assessment Exercise”(IAE). To carry out the IEA, the Council has normally issued detailed guidelines in advance. So, in April 1999, the BBSRC announced that the second IEA would be conducted in 2001, following the first IAE carried out in 1996-1997, and

published the IAE 2001 guideline for the institutes to be evaluated.

According to this guideline, the 2001 IAE comprised the following four components : (1) assessments by expert panels of the quality of science (QS), (2) an assessment by a panel of knowledge transfer (KT), (3) an assessment by the BBSRC office of competitive funding (CF) and (4) an assessment of the institute as a whole by a visiting group (VG), which also include an assessment of that institute's strategic relevance including the advanced training provided. This guideline deals with very detailed contents for each component such as the scope of assessment, the assessment criteria, the assessment ratings, the assessment process, the information requirements and so on. In its annex, it also contains very detailed and specified sheets for each component which institutes are required to submit to the Council. The outcome of this exercise informs the Council's allocation of the CSG to the BBSRC-sponsored institutes and Horticulture Research International for the four financial years from 2002-2003 to 2005-2006.

#### **4. 3. Recent Evaluation Exercises in S&T and R&D Sectors**

In the UK, according to the basic guidance of both the central governments and the executive departments, diverse evaluation activities have been conducted at the different levels of the S&T and R&D sectors. As already indicated, these evaluations are grouped into three types : the evaluation of science and research policy, the evaluation of institutions (e.g. research councils and RCs institutes) and the evaluation of S&T and R&D programmes. Consequently, in the UK, five different types of evaluation have normally been carried out in the S&T and R&D sectors. Recent evaluation exercises are summarised in <Table 4-6>.

In particular, as at April 2002, some evaluations of science and research policy and research councils have been completed and their findings have also communicated to the general public, while some are now in progress and will be completed in the near future. So, I have chosen completed evaluations as examples for case studies. In fact, I have selected the cross-cutting review of science and research by HM Treasury (2000) and the quinquennial review of the six grant awarding research councils by the OST (2001) as marked (\*) in <Table 4-6>, as the analytical examples.

In what follows, I shall analyse and describe five evaluation exercises on the basis of four main components and 11 sub-components of the meta-evaluation model suggested in Chapter 3. For this, in each case, the guidelines concerned, evaluation reports or review papers, the related documents in the websites and the results of interviews were used as basic materials for analysis

**<Table 4-6> Recent Evaluation Exercises in S&T and R&D Sectors (UK)**

Type of Evaluation		Recent Evaluation Exercises
<b>Evaluation of Science and Research Policy</b>		<ul style="list-style-type: none"> <li>- Cross-cutting Review of Science and Research (2000) (*)</li> <li>- Cross-cutting Review of Science and Research (2002)</li> </ul>
<b>Evaluation of Research Institution</b>	<b>Evaluation of Research council</b>	<ul style="list-style-type: none"> <li>- Quinquennial Review of the Six Grant Awarding Research Councils (2001) (*)</li> <li>- Quinquennial Review of the CCLRC (2002)</li> </ul>
	<b>Evaluation of Research Institute</b>	<ul style="list-style-type: none"> <li>- Institute Assessment Exercise (IAE) by the BBSRC (2001)</li> </ul>
<b>Evaluation of S&amp;T / R&amp;D Programme</b>	<b>Evaluation of S&amp;T Programme</b>	<ul style="list-style-type: none"> <li>- Evaluation of Funding for UK Civil Space Activity by the DTI (2001)</li> </ul>
	<b>Evaluation of R&amp;D Programme</b>	<ul style="list-style-type: none"> <li>- Review of the Gastrointestinal Response to Food Programme by the BBSRC (2000)</li> </ul>

#### ***4.3.1. Evaluation of Science and Research Policy : Cross-cutting Review of Science and Research by HM Treasury (2000)***

From late 1999 to early 2000, HM Treasury conducted the first cross-cutting review of science and research in advance of the Spending Review 2000 (SR 2000). The review was undertaken in order for government departments concerned with the science and technology programmes and policies to develop co-operation with one another and make a variety of new working arrangements. This review was also one of the 15 full cross-departmental studies undertaken as a part of SR 2000. Most outcomes of this cross-cutting review have been adopted as crucial policy means and alternatives in the SR 2000 White Paper as well as in the Science and Innovation White Paper 2000, *Excellence and Opportunity : A Science and Innovation Policy for the 21st Century*, produced by the DTI in July 2000.

#### ***4.3.1.1. Evaluation Paradigm***

##### **(1) Evaluation Purpose**

The remit of this review was to ensure that science, which has been supported from public funds, is of maximum excellence and that it is exploited to the benefit of the economy at large. The review considered the contribution of university research to economic growth and productivity. It examined how well current funding levels and mechanisms for publicly funded science and research maximised this wider economic contribution. In particular, the review was asked to (1) assess how well current funding mechanisms promote research excellence and the optimal balance between types and disciplines of research, (2) develop proposals for funding university capital requirement which make the money follow and bolster excellent research, and encourage universities to make proper provision for maintenance, renewal and new investment, (3) make proposals to ensure that exploitation of research and the interface with business are embedded in the core mission of universities and (4) ensure that government departmental science and technology spending plans sit well with funding proposals for the science and engineering base, and that they are underpinned strategically (HM Treasury, 2000a : 2).

From this remit, we can draw some major evaluation purposes of the 2000 cross-cutting review. In this review, as in to those of most policy evaluations, both accountability and policy/programme improvement were announced explicitly as two main evaluation purposes (Goldenberg, 1983 ; Arvidsson, 1986 : 627 ; Lee and Sampson, 1990 : 157 ; Hudson, Mayne and Thomlison, 1992 : 5). For the former purpose, the review examined whether science supported from public funds has been of maximum excellence and to what extent it has contributed to economic growth and productivity. Meanwhile, for the latter purpose, it also examined how well current funding levels and mechanisms for publicly funded science and research have maximised this wider economic contribution.

##### **(2) Evaluation Object**

The review looked at science and engineering research supported by public funds through the Higher Education Funding Council for England (HEFCE) and research councils as well as funding by other government departments of research and development in universities and public sector research establishments (PSREs). As a result, it did not cover the devolved administrations in Scotland, Wales and Northern Ireland. It therefore

made recommendations in relation to the HEFCE, the research councils that cover the whole of the UK and the science and technology activities of UK government departments (HM Treasury, 2000a : 2).

Even though this review was initially intended to investigate the cross-departmental science and research activity as a whole, it appears that the level of the evaluation object included both science and research “policy” in civil departments and their main S&T and R&D “programmes”. On the other hand, the scope of the evaluation object was, in principle, confined to science and research activities in England conducted, through the involvement of the HEFCE and the research councils. However, in practise, the review considered funding of the SET base, which inevitably included consideration of the contribution of other funders of research in universities, such as charities, industry, the EU and government departments. Consequently, it seems that the scope of the evaluation object dealt with all SET activities in the UK as a whole.

### **(3) Evaluation Type**

The evaluation type has not been clearly announced in this review. However, from the analysis of remits of the review (HM Treasury, 2000a : 2), we know that it implicitly intended to carry out an “output” evaluation and, where appropriate, an “impact” evaluation. Namely, in order to judge the excellence of publicly funded science and its contribution to the economic growth and productivity, it needed to carry out an “output” evaluation to measure direct and short term results of the science and research policy concerned and an “impact” evaluation to investigate both the direct and short term as well as indirect and long term influences of that policy on the economy and society as a whole.

#### ***4.3.1.2. Evaluation Resources***

##### **(1) Evaluation Manpower**

For this review, the staff of the Enterprise Team in HM Treasury mainly prepared the evaluation plan and implemented the evaluation work. As at April 2002, there were three officials in the Enterprise Team, although only one member was involved in this review. Meanwhile, we can see that the staff of the Enterprise Team have carried out many other missions as well as evaluation work. In addition, external advisers and experts who represented other related departments such as the DTI and the DfES, the university and



private sectors were also involved in this review for a short period (HM Treasury, 2000a : 3). The total external manpower for this evaluation was five.

As a result, it is apparent that the 2000 cross-cutting review of science and research was conducted through collaborative work between the staff of the Enterprise Team in HM Treasury and external advisers and experts. However, while the staff of the Enterprise Team carried out this evaluation as their permanent work, the external manpower did this as temporary work for a period of some months.

## **(2) Evaluation Unit**

In HM Treasury, there are 19 spending teams that oversee the work of each department in order to monitor their spending on their missions. Most of the spending teams are in the Public Services Directorate (PSD), but some are in other directorates closer to the policy issues. The spending teams are the first point of contact for people in spending departments concerned with financial planning and control. They are organised so that each department only has to deal with the people in one spending team for most day-to-day business (Cabinet Office, 2002a : 2).

In the area of science and research, the Enterprise Team acts as the spending team for the DTI and has the lead policy responsibility within HM Treasury for science, innovation and enterprise. The Enterprise Team influences policy for science and innovation by, for example, (1) ensuring that public funding concentrates on excellence and rewards exploitation, (2) encouraging universities and public sector research establishments to increase their ability to manage and exploit intellectual property effectively and (3) encouraging more innovation in business. This team sits within the Enterprise and Growth Unit in the Finance, Regulation and Industry Directorate (FRID) (HM Treasury, 2001a : 3).

Consequently, the Enterprise Team in HM Treasury played a role as an evaluation unit in the conduct of the 2000 cross-cutting review of science and research, a special type of evaluation on science and research policy at nation level.

## **(3) Evaluation Information**

In this review, important information was drawn from the two kinds of sources. One was

the existing research which had been carried out to investigate the connections between science and economic growth. The followings are examples of existing research analysed for this review :

- Talent, not Technology : Publicly Funded Research and Innovation in the UK (SPRU, The University of Sussex : 2000)
- Technology Transfer : The US Experience (Report of a CVCP mission funded by the Gatsby Charitable Foundation : 1999)
- University Staff and the Knowledge Based Economy (The Association of University Teachers and the University of London Institute of Education : 2000)

The other source was the extensive consultation which was conducted with senior university academics, the Committee of Vice Chancellors and Principals (CVCP), the Higher Education Funding Council for England (HEFCE) and the Wellcome Trust (HM Treasury, 2000a : 3).

Therefore, in the 2000 cross-cutting review, much more use was made of qualitative, sometimes slightly subjective, information than of quantitative one in conducting the evaluation of science and research policy.

#### ***4.3.1.3. Evaluation Performance***

##### **(1) Evaluator**

This review was led by the Minister for Science and Innovation in the DTI and he was supported by the “Steering Group”. The steering group was chaired by a senior HM Treasury official and included representatives from the DTI and the DfES and external advisers from the universities and private sectors. The detailed roles and missions of the steering group were to conduct an actual evaluation and to make many recommendations for future development of UK SET activities (HM Treasury, 2000a : 3 ; HM Treasury, 2001a : 3).

Thus, in this review, it seems that while both internal HM Treasury evaluators and external evaluators were involved, external evaluators, including experts from universities and the private sector, played more important roles than internal ones did.

## **(2) Evaluation Method**

The steering group reviewed the available research literature on the connections between science and economic growth. It also conducted an extensive consultation, seeking the views of more than a hundred senior academics in 12 universities, as well as those of the Committee of Vice Chancellors and Principals (CVCP), the Higher Education Funding Council for England (HEFCE) and the Wellcome Trust (HM Treasury, 2000a : 3).

Hence, in this review, a qualitative method, including a literature review and extensive consultations, was adopted as the main evaluation method.

## **(3) Evaluation Indicator**

In the 2000 cross-cutting review, in order to judge the excellence and the economic contribution of science and research supported by public sector, the steering group, namely the evaluation team, adopted the followings as major evaluation indicators : (HM Treasury, 2000a : 7, 13, 22).

- producing people with knowledge : trained scientists and engineers who generate knowledge in the science base and apply their expertise to developing new products and processes for industry
- generating and disseminating new knowledge : scientific and technological developments with applications which can lead to new products and processes.
- degree of concentration on RAE funding on the best facilities : because the RAE is the key driver for excellence in the funding of research.
- commercialisation and collaboration with industry : this indicator was further divided into the following three sub-indicators :
  - capacity building : for example, recruiting experts in technology transfer, building a proper interface with business and establishing robust systems for identifying and managing intellectual property portfolios etc.
  - seed funding for early stage feasibility work on commercialisation proposals
  - services for industry for which the costs cannot be fully recovered (for example, the provision of advice and expertise for SMEs and research collaborations involving several firms)

Thus, in this evaluation, both qualitative and quantitative indicators were simultaneously adopted, even though the usage ratio between them and their weights were different.

#### **4.3.1.4. Evaluation Utilisation**

##### **(1) Evaluation Report**

The final report of the 2000 crossing-cutting review was published and can be found on the website of HM Treasury ([http://www.hm-treasury.gov.uk/sr2000/associated/science\\_231100.pdf](http://www.hm-treasury.gov.uk/sr2000/associated/science_231100.pdf)). This report contains many evaluation components which have been thought of as main factors such as an evaluation purpose, an evaluation object, the evaluators, the evaluation methods, main recommendations and evaluation utilisation. Its core contents were also summarised in “Section V : Cross-Departmental Review” in the *Spending Review 2000 White Paper*.

##### **(2) Utilisation Type**

This review made a number of recommendations which fell into five categories, namely (1) public funding for science, (2) dual support : the funding system for university research, (3) capital funding for research infrastructure, (4) commercialisation of research and collaboration with industry (knowledge transfer) and (5) departmental spending on science and technology. As a result, the thrust of the review and many of the detailed recommendations were reflected in the SR 2000 settlement and the Science and Innovation White Paper 2000. In accordance with the conclusion and recommendation of the review, the *Spending Review 2000 White Paper* and the Science and Innovation White Paper 2000 confirmed that the government would increase the combined DTI and DfES spending on science and research by an annual average of 5.4% in real terms from 2000-2001 until 2003-2004 (HM Treasury, 2000a : 3-6 ; HM Treasury, 2001a : 3).

Key outcomes reflected in the two White Papers were (1) a new £1 billion capital programme in partnership with the Wellcome Trust to renew the infrastructure for science, providing world class buildings and equipment for leading edge research, (2) a £250 million boost to research in the following key new areas that will shape life in the 21st century : genomics, e-science and basic technology such as nanotechnology, quantum computing and bio-engineering, (3) additional funding to increase over three years the basic support for post-graduate research students to £9,000 a year, (4) a Higher Education Innovation Fund of £140 million over three years (incorporating the Higher Education Research OUT to Business and the Community fund) to build on universities' potential as drivers of growth in the knowledge economy. This will triple existing funding by

2003-2004, increasing universities' capability to work with industry, and particularly with small firms, (5) one further round of the University Challenge competition, to provide seed venture funding for knowledge transfer, and £15 million more into Science and Enterprise Centres to bring business skills into the science curriculum, (6) a commitment to publish science and innovation strategies for government departments, with anticipated real term rises in the biggest civil R&D programmes, which are those of the MAFF (now the DEFRA), the DoH and the DETR (now the DfT) and (7) a commitment to investigate whether there are any structural causes of under investment in universities' research capital (HM Treasury, 2000a : 6 ; DTI, 2000b).

Therefore, it can be said that the findings of this review were fully utilised for both short term and long term policy. While inputs to the SR 2000 settlement were short term policy means for R&D programme and science and research policy, those to the Science and Innovation White Paper 2000 were seen as long-term policy means in the area of S&T and R&D policy.

#### ***4.3.2. Evaluation of Research Councils : Quinquennial Review of the Six Grant Awarding Research Councils by the DTI and the OST (2001)***

Between February and November 2001, both the DTI and the OST conducted the quinquennial review of the six grant awarding research councils, under the terms of the Cabinet Office guidance document : *How to Review Agencies and Non-Departmental Public Bodies to Improve the Quality and Effectiveness of Public Services*, published on 31 January 2000. This was the first quinquennial review of the research councils since their re-establishment in 1994 as a result of the Science and Technology White Paper 1993, *Realising our Potential : A Strategy for Science, Engineering and Technology*. The quinquennial review was conducted in two stages. While the first stage judged the roles and organisation of the research councils, by reference to their Charters and missions and evidence of past practice, the second stage examined the efficiency and effectiveness of their operations. The reports of stage one and stage two were published on 19 July 2001 and 4 December 2001, respectively. The outcomes will be one of the most important inputs to the 2002 cross-cutting review of science and research, a part of Spending Review 2002, as well as the future policy development in relation to the research

councils' management.

Meanwhile, the Council for the Central Laboratory of the Research Councils (CCLRC) has been subject to a separate quinquennial review. The reports of the stage one and two of the review were published on 26 February 2001 and 30 April 2002, respectively (DTI/OST, 2001a ; DTI/OST, 2002).

#### ***4.3.2.1. Evaluation Paradigm***

##### **(1) Evaluation Purpose**

Under the Cabinet Office's 2000 guidance document, this review was conducted in two stages whose purposes were slightly different. The purpose of the first stage was to examine "the role and organisational type" of the research councils, by reference to their Charters and missions and evidence of past practice. In this stage, it was also intended to set the detailed terms of reference for the second stage. Meanwhile, the purpose of the second stage was to judge "the efficiency and effectiveness" of the Councils' operations and look for further opportunities for improving performance. Thus, stage two of the review investigated four broad themes, which were identified following the extensive stage one consultation. These were : (1) mission, structure and governance, (2) relations between the research councils and their stakeholders, (3) methods of priority setting and decision-making and (4) the Councils' management and internal processes (DTI/OST, 2001c : 11).

Therefore, it appears that the main remit of the quinquennial review of the six research councils was to examine their past performances and to judge the appropriateness of their operational status. An additional remit, on the basis of these outcomes, was to suggest future improvements to the management of research councils. As a result, in the quinquennial review in 2001, the government intended to secure two direct evaluation purposes : "accountability" and the "improvement" of management, which have usually been thought of as a main purpose of policy evaluation in general. In addition to these, we may think that the "knowledge enhancement" for the R&D management system in the public sector was also one of the evaluation purposes. However, this purpose was not expressed in an explicit and direct way in either the review plan or the evaluation report.

## **(2) Evaluation Object**

From the evaluation purposes, we may see that the evaluation object of quinquennial review covered both past performance since 1994 and the future opportunities for improvement of the research councils' management. However, in view of the challenge of reviewing six organisations within a reasonable timescale, the review team concentrated on the systems and processes used by the research councils and the OST rather than the content of their policies and strategies. They also took a forward-looking view, carrying out retrospective analysis where it was necessary to do this, and considered in particular the effectiveness and efficiency of the research councils as a group (DTI/OST, 2001c : 12).

Hence, it was apparent that while the scope of the evaluation object included a retrospective examination of past performance as well as a forward-looking review of the future management plan in the research councils, the level of the evaluation object was confined to the operational aspect, rather than extending to the strategies and high-level policies.

## **(3) Evaluation Type**

To carry out the evaluation purpose mentioned above, both the "output" evaluation and "assessment" of future management were given emphasis. That is, the retrospective examination of past performance is the typical type of output evaluation, although the forward looking review of future management plans in the research councils could be considered as assessment tasks in R&D management.

### ***4.3.2.2. Evaluation Resources***

#### **(1) Evaluation Manpower**

In order to engage as wide a range of stakeholders as possible in the review, three groups of evaluation manpower with different missions were established for a temporary period. The review was overseen by a "steering group" chaired by the Director General of Research Councils (DGRC). The steering group had 17 members including the chairman, who comprised six councils' Chief Executives, six members from academia and two members from business. The actual review was conducted, under the guidance of a steering group, by an independent "review team" led by a senior DTI official and

included five members with one from each of the two related government departments, research council, research charity and business. In the second stage of the review, in particular, four “working groups” were set up to examine the four themes. A member of the steering group chaired each of the working groups. The working groups had a wide membership drawn from academia, business, research charities, government and elsewhere, and included representatives from the research councils. A detailed breakdown of the members in each working group according to institutions to which they belonged is given in <Table 4-7>. They were supported by secretariats which included both members of the review team and research council staff (DTI/OST, 2001c ; 13).

**<Table 4-7> Breakdown of the Working Group Members in Quinquennial Review**

Working Group	Chair	Members									Total	Secretariat
		A	B	G	RC	FC	C	RI	P	E		
<b>A</b>	1(A)	3	1	4	6	1	1	-	-	-	17	3
<b>B</b>	1(A)	4	2	2	6	1	1	3	-	-	20	4
<b>C</b>	1(RB)	5	3	7	7	-	-	3	1	2	29	4
<b>D</b>	1(A)	5	1	2	6	-	3	1	-	-	19	4
<b>Total</b>	4	17	7	15	25	2	5	7	1	2	85	15

A : academia

RC : research councils

RI : research institutes

E : elsewhere

B : business

FC : funding councils

RB : research board

G : government

C : research charities

P : parliament

## (2) Evaluation Unit

Even though a lot of organisations and external experts were involved in this review, the “Science and Engineering Base Group (SEBG)” in the OST,<sup>5</sup> led by the DGRC, acted as

<sup>5</sup> The OST comprises two main support groups. One is the “Trans-departmental Science and Technology Group” (TDSTG) supporting the Chief Scientific Adviser (CSA), the other is the “Science and Engineering Base Group” (SEBG) supporting the DGRC.

Of these, the SEBG aims to ensure Science Budget funded programmes, together with the DfES’s support for the Science and Engineering Base (SEB), meet the country’s future requirements, taking account of Foresight and other expressions of user needs. Related issues include (1) the quality, number and subject balance of scientists and engineers qualified at postgraduate and post-doctoral levels, (2) fundamental and strategic research and the maintenance of the scientific and technical knowledge base and (3) access to advantages in science, engineering and technology being generated outside the UK.

In addressing these, the SEBG has missions to ensure that (1) the research councils and other funded bodies work efficiently and effectively, (2) there is an increase in the number of female scientists and engineers and that women are properly represented in senior positions in public bodies and SET based firms,



the evaluation unit. As one of the roles of the SEBG is to ensure that the research councils and other publicly funded bodies work efficiently and effectively (OST, 2002), so the SEBG had to play the role of both secretariat and evaluation unit in the quinquennial review on the research councils.

Meanwhile, a new mini-directorate was set up to actually conduct the review, although there were three following directorates within the SEBG : Research Councils Directorate, Exploitation Directorate, and Finance, Policy and Cooperate Affairs Directorate (OST, 2000a). This mini-directorate was temporary and closed when the review had been completed. It was also independent of the rest of the OST. At the time the review was being carried out, there were a core group of around six people in the directorate.

### **(3) Evaluation Information**

In this review, most of the important information came from three kinds of sources : (1) an extensive consultation with key stakeholders, (2) related literature and existing studies and (3) an in-depth discussion with staff from the research councils. In the course of the review, the two public consultations conducted in each stage by the review team provided a lot of useful information to that team and working groups. Additionally Working Group C's independent consultation in stage two was the source of further invaluable information for the review team and the other working groups (DTI/OST, 2001c : 12-14). A more detailed explanation of this will be given in the following section on the "evaluation method".

Consequently, diverse qualitative information was utilised as the main source for this review. However, it seems that rather infrequent use may have been made of quantitative information.

#### ***4.3.2.3. Evaluation Performance***

##### **(1) Evaluator**

---

(3) best use is made of all available SET expertise in the membership of relevant public bodies and committees, (4) the connections between academics, industrialists and others in the SEB are improved so as to promote sustainable growth and preserve and improve the quality of life, (5) there is an improvement in the public understanding and appreciation of the SET and (6) appropriate statistics showing the state of the UK SET are collected, analysed and disseminated (OST, 2002).

As already described, three groups of evaluation manpower, namely steering group, review team and four working groups, were involved in this review. Of these, it was a few officials of the review team from the DTI and the OST who were actually responsible for the evaluation work. On the other hand, external advisers from both academia and the related government departments numbered 45 out of the total of 85 members of the four working groups. So, we may see that this review was conducted by the external evaluators, rather than by the internal ones.

## **(2) Evaluation Method**

For the quinquennial review, three kinds of evaluation method were adopted : (1) public consultation with key stakeholders, (2) review of the related literature and studies and (3) in depth discussion with the members of staff in research councils.

In each stage of this review, both the review team and the working group undertook a substantial consultation exercise involving key stakeholders. Through this work, they analysed the comments made during the consultations and took account of these in their consideration of the issues concerned. In the first stage, in particular, the consultation on the 11 issues <sup>6</sup> in relation to the remit of the review were conducted in advance of the formal commencement of stage one of the review on the 13th February 2001. The review team sought to publicise the existence of the review as widely as possible in order to gain the views of a wide range of stakeholders. The main vehicle for this consultation was a letter in late October 2000, inviting written submissions, which was sent to a wide range of stakeholders, posted on the OST website, and publicised through a DTI press release.

The review team also contacted a number of relevant organisations and professional bodies and requested that their members be advised of the review. The closing date for

---

<sup>6</sup> In stage one, the 11 issues on which consultation was made were as follows (DTI/OST, 2001b) :

- (1) The current scope of research council missions, boundaries, and cross-council collaboration
- (2) The relationship between the research councils and the OST
- (3) research council membership
- (4) The balance between science funding and administrative overheads
- (5) The extent to which strategic and basic research needs are being addressed
- (6) The effectiveness of engagement with academic and user communities
- (7) The effectiveness of exploitation and knowledge transfer (including trained people)
- (8) The standards of service to customers and the opportunities/efficiencies from the wider adoption of information technology (IT)
- (9) Transparency and openness of operations
- (10) Relationships with research council owned institutes and other directly funded units

written submissions to stage one of the review was 15th March 2001. As a result, a total of 113 submissions were received. The stakeholder groups who provided submissions and the number of submissions they provided are shown in <Table 4-8> (DTI/OST, 2001b : 3).

**<Table 4-8> Organisations of Submission in Stage One of Quinquennial Review**

<b>Organisation</b>	<b>Submission Number (%)</b>
CST (Councils for Science and Technology)	1 (0.9%)
Research Councils and individual RC members	19 (16.8%)
Universities	45 (39.8%)
Public Sector Research Establishments (PSREs)	6 (5.3%)
Funding Councils and Agencies	5 (4.4%)
Government Departments	4 (3.6%)
Learned Societies and Professional Bodies	9 (8.0%)
Representative Organisations	11 (9.7%)
Individual Submissions	13 (11.5%)

**<Table 4-9> Breakdown of the Respondents in Stage Two of Quinquennial Review**

<b>Respondents</b>	<b>Submission Number (%)</b>
Universities (institutional, departmental and individual academics)	105 (49.1%)
Learned Societies & Professional Institutions	32 (14.9%)
Business (companies, trade unions, individuals)	22 (10.3%)
Government Departments and other Public Sector Bodies	25 (11.7%)
Research Institutes and Research Associations (including individuals)	13 (6.1%)
Charities	6 (2.8%)
Staff Representative Organisations	6 (2.8%)
Other Individuals	5 (2.3%)

There was also a written consultation for stage two of the review. It was announced on April 2001 and published on the OST website with a final response date of 30 June 2001. The consultation letter invited comments on the four issues which formed the basis of the terms of reference of the four working groups. It was also circulated widely to key organisations, including industry representative organisations, government departments and devolved administrations, universities, regional development agencies, select

---

(11) Promotion of the public understanding of science

committees, funding agencies including research charities, learned societies and professional institutions. The review team invited the research councils to distribute the consultation letter by e-mail to their members, committees and panel members, and also to grant holders, research institutes, user communities and staff unions, and to post the letter on their website. The total number of responses was 214, and the breakdown of the respondents according to sector is shown in <Table 4-9>. All the responses, except those that were marked confidential, were circulated to the research councils, so that they could make use of them in the course of the review, and utilise them as a useful resource in the future (DTI/OST, 2001c : 14, 103).

Between its two meetings, Working Group C also undertook a further written and oral consultation with a range of research council stakeholders. As part of its special consideration of the relationship between the research councils and their main stakeholder groups, it invited the research councils to (1) identify their main categories of stakeholder, (2) summarise the mechanisms and processes that they used to interact with their stakeholders, (3) name some selected stakeholders whom Working Group C could consult and (4) give their impressions of how effectively the interface worked. Members of the Working Group C and the co-secretaries consulted stakeholders within these categories to gauge their views on the quality of their interactions with the research councils. For this consultation, over 120 trade associations and selected industrial representatives were contacted by e-mail. Also, established networks of the University Research Administrators Group Network (RAGnet) were used to gather views from its 400 research administrators in research centres and departments in HEIs. A range of contacts in government departments, devolved administrations and Parliament were contacted personally by members of the Working Group C for their perceptions of the research councils. As a result, total respondents to this stakeholder consultation were 23 organisations, of which 10 were from the public sector and 13 from the business sector. All the views collected in these ways were collated with relevant comments provided through the stage two written consultations and fed back to working groups. They were made available to the research councils and were used by working groups to help formulate its conclusions and recommendations (DTI/OST, 2001c : 14, 112-113).

Next, a lot of literature was utilised for this review. A number of related reviews were being carried out or had been commissioned in parallel with this review. So, the review

team took account of these works but did not duplicate them in the review. The following are some examples of the literature used in the review (DTI/OST, 2001c : 11-12) :

- The 2002 Spending Review, of which the cross-cutting review of science and research is a part
- The 'Transparency Review' of university research and its follow up
- The OST/HMT/DfES study of under investment in research infrastructure
- The Funding Councils' review of research policy and funding
- An OST review of work on science communication
- The independent review of the supply of scientists and engineers announced at the time of the 2001 budget (the 'Robert Report')
- Are we realising our potential? (The Sixth Report of the House of Commons Select Committee on Science and Technology, April 2001)

Also, the review team conducted in-depth interviews and regular discussions with the research councils' staff concerned. They reported that they had conducted the review in an open and participative fashion, involving the research councils fully in all aspects of the work. They consulted regularly with nominated review co-ordinators in each research council. They also had meetings with the research council staff representative associations and town meetings with staff. So, the research councils worked together during the course of the review in a way which exemplified the approach which the review team was encouraging for the future (DTI/OST, 2001c : 13-14).

Therefore, it appears that, in the quinquennial review in 2001, qualitative analytical approaches including public consultation, literature review and in-depth interviews were adopted as the main evaluation method.

### **(3) Evaluation Criteria and Indicator**

In this review, the evaluation criteria and indicators were not clearly described. However, we can infer these, in particular the evaluation criteria, from the evaluation purpose. As already noted, the evaluation purposes of the review were to examine the role and organisational type of the research councils in the first stage, and to judge the efficiency and effectiveness of the research councils' operations in the second stage (DTI/OST, 2001b ; 2001c). Thus, we may consider the following three factors as principal evaluation criteria : the "appropriateness of the research councils' organisational status", the

“efficiency of the research councils’ management” and the “effectiveness of the research councils’ management”. Therefore, it is apparent that both qualitative and quantitative criteria were utilised, even though the ratio of their usage was slightly different in the actual evaluation process.

Meanwhile, there was no clear identification of the evaluation indicators in the related official documents including the two review reports.

#### ***4.3.2.4. Evaluation Utilisation***

##### **(1) Evaluation Report**

The reports of stage one and stage two of the review were published on 19 July 2001 and 4 December 2001, respectively. They have also been available on the OST website (<http://www.dti.gov.uk/ost/whatsnew/index.htm>). These reports contain a lot of evaluation components which have been thought of as main factors in policy evaluation, such as the evaluation purpose, the evaluation object, the evaluators, the evaluation methods, principal recommendations and evaluation utilisation. In particular, some parts of the review report in stage two were devoted to the description of the evaluation methodology and the sources of the evaluation information (DTI/OST, 2002b : 11-14).

The report of stage two of the review, which is regarded as a final report of this quinquennial review, comprised four main parts : contents, findings, annexes and appendices. Of these, the findings section dealt with seven issues <sup>7</sup> on the future research councils’ development and made many recommendations regarding these (DTI/OST, 2002c).

---

<sup>7</sup> The following are the findings from the review of stage two in summary form (DTI/OST, 2002b) :

- (1) Collective leadership and the research councils
- (2) Priorities between the UK and international programmes
- (3) Allocation of research council funds :
  - Decision making process
  - Research councils’ institutes : the research councils as providers of research
  - Postgraduate training
  - Knowledge transfer
- (4) Relationship with stakeholders, including peer review and communicating science
- (5) Convergence of research council’s management process
- (6) Monitoring
- (7) Resources and implementation

## **(2) Utilisation Type**

The results of this review will be one of the most important inputs to the 2002 cross-cutting review of science and research, a part of the Spending Review 2002, whose contents are expected to announce in summer 2002, as well as the future policy development in relation to the research councils.

As already described, stage one of the review focused on the identification of issues for detailed consideration in stage two, together with an analysis of the effectiveness of the continuation of the current status of the research councils as NDPBs under Royal Charter. The principal recommendation of stage one was that the research councils should continue to be executive NDPBs (DTI/OST, 2002b : 19).

Following the outcome of the review of stage one, the review of stage two focused on the four themes already noted above. The major conclusions and recommendations included the followings (DTI/OST, 2002c : 5-10) :

- (1) The establishment of a new strategy group comprising the research council chief executives and the DGRC, to achieve a more unified approach, and to provide a framework for cross-council working at all levels.
- (2) The development of a clearer identity and mission, whereby the research councils will be able to establish stronger links with the other major science funding organisations, including the funding councils, government departments and major charities.
- (3) A closer relationship between the research councils and other key stakeholders, including the universities and the business and public service organisations which use their research and expertise.

As of February 2002, the findings of the 2001 quinquennial review had not yet been utilised as concrete policy means. However, in the near future, they will be a very important input to the Spending Review 2002 White Paper. Also, just as this review recommended, in two years' time, the OST should evaluate whether these findings have been achieved (DTI/OST, 2001c : 4), so it is expected that the related government departments, including HM Treasury, the DTI, the DfES and the OST, will adopt these recommendations.

Therefore, it seems that the evaluation findings will be fully utilised to improve the

research councils' management systems as well as being used in the broader perspective to re-arrange their roles and missions in line with UK science and technology policy.

#### ***4.3.3. Evaluation of Research Institutes : The Institute Assessment Exercise by the BBSRC (2001)***

Following the first Institute Assessment Exercise (IAE) in 1996-1997, the BBSRC conducted the second IAE between April 2001 and February 2002. It circulated its guidelines for this work, *IAE 2001 Guideline*, to its staff and the research institutes concerned in April 1999 and again in April 2000 in advance of the IAE 2001. As already noted, the BBSRC supports research and training in over 60 universities and in eight sponsored research institutes. Some 35% of the BBSRC funds have been allocated to its institutes, which deliver fundamental and strategic research primarily in the non-medical life sciences (BBSRC, 2001a).

##### ***4.3.3.1. Evaluation Paradigm***

###### **(1) Evaluation Purpose**

The IAE 2001 was conducted with the intention of informing the eight sponsored institutes and Horticulture Research International (HRI) of the Council's allocation of the Core Strategic Grant (CSG) for the four financial years from 2002-2003 to 2005-2006 (BBSRC, 1999a : 1).

Meanwhile, as the public sector sponsor of eight institutes employing 3,500 BBSRC staff, with a turnover of £140 million and property assets of £300 million, the BBSRC is responsible, together with the governing bodies, for ensuring high standards of accountability, propriety and the efficient use of resources (BBSRC, 1999b : 23). To achieve these, the IAE has been considered as the most important among a lot of policy means.<sup>8</sup>

---

<sup>8</sup> In order to achieve high standards of performance, the BBSRC has : (1) made clear statements of responsibility agreed between the BBSRC and each institute, (2) approved and monitored annual business and efficiency plans, (3) operated a robust and transparent research management accounting system, (4) determined pay and conditions for institute staff as the BBSRC employees, (5) used performance indicators and conducted major reviews of each institute every four years and (6) encouraged institutes to work together and share resources (BBSRC, 1999b : 23).



From these statements, it is apparent that the BBSRC conducted the IAE 2001 in order to secure both the “accountability” and “improvements” of the expenditures of research fund.

## (2) Evaluation Object

The IAE 2001 comprised the following retrospective elements which related to the five year period from April 1996 to March 2001 : (1) assessments by expert panels of the quality of science (QS), (2) an assessment by a panel of knowledge transfer (KT), (3) an assessment by the BBSRC Office of competitive funding (CF) and (4) an assessment of the institute as a whole by a visiting group (VG) (BBSRC, 1999a : 1).

<Table 4-10> Evaluation Object of the IAE 2001

Evaluation Object	Evaluation Sub-object	Detailed Evaluation Items
<b>Quality of Science</b>	Publication	<ul style="list-style-type: none"> <li>- refereed articles</li> <li>- edited books and chapters</li> <li>- conference proceedings and abstracts</li> </ul>
	Measures of Esteem	<ul style="list-style-type: none"> <li>- prizes</li> <li>- other measures of esteem <ul style="list-style-type: none"> <li>• invitations to give key-note addresses at international conferences</li> <li>• chair/member of peer review panels</li> <li>• chair/member of advisory/policy group</li> <li>• chair/member of editorial boards of leading refereed journals</li> </ul> </li> </ul>
<b>Knowledge Transfer</b>	Retrospective Assessment (last five years)	<ul style="list-style-type: none"> <li>- exploitation of institute research through licensing and company start-up formation</li> <li>- knowledge transfer to both industrial and policy users</li> <li>- institute involvement in research and training in collaboration with users</li> <li>- staff involvement in exchanges and consultancies</li> <li>- open days, workshop and site visits etc.</li> </ul>
	Prospective Assessment (next five years)	<ul style="list-style-type: none"> <li>- the adequacy of proposed arrangements to undertake KT</li> <li>- the influence of these arrangements likely to have on the effectiveness of future performance in KT</li> </ul>
<b>Competitive Funding</b>	-	<ul style="list-style-type: none"> <li>- the competitively won external income</li> </ul>
<b>Visiting Group Assessment</b>	The Totality of Research	<ul style="list-style-type: none"> <li>- the progress made in implementing the recommendations of the previous VG</li> <li>- an integrated assessment of institute's performance</li> </ul>
	Forward Research Plans	<ul style="list-style-type: none"> <li>- its strategic relevance, including an assessment of the advanced research training</li> </ul>

On the basis of these, we may infer that there were four evaluation objects : the quality of science, knowledge transfer, competitive funding and visiting group assessment. Each evaluation object was again divided into more detailed evaluation sub-objects, which consisted of specified evaluation items, as shown in <Table 4-10>. While some objects such as knowledge transfer and visiting group assessment covered both past performance and future research plans of individual institutes to be evaluated, other objects such as science quality and competitive funding included only past performance.

### **(3) Evaluation Type**

Regarding evaluation type, there has not been an explicit announcement within the official documents including the IAE 2001 guideline. However, it could be drawn from the evaluation object above mentioned. Consequently, in this evaluation, an “output” evaluation and “assessment” of the research plan in each institute were conducted simultaneously. While the measurement of past performance in relation to science quality, knowledge transfer and competitive funding was a kind of output evaluation, the prospective investigation of knowledge transfer and of forward research plans was typical assessment task in R&D management process as a whole.

#### ***4.3.3.2. Evaluation Resources***

##### **(1) Evaluation Manpower**

For the IAE 2001, two groups of evaluation manpower were involved. While one was the external expert panels including visiting groups, the other was the BBSRC's officers. The six panels, each of 12 members from academia, business and government, carried out the evaluation on the science quality on the basis of both formulae published in advance and information given by each institute. The panels covered the following six areas : (1) agri-food, (2) animal sciences, (3) engineering and biological systems, (4) genes and developmental biology, (5) molecular sciences, biochemistry and cell biology and (6) plant and microbial sciences. These panels included assessors from the 1996 and 2001 Research Assessment Exercises (RAEs) and the BBSRC's research committee and Network Group members, supplemented as necessary to cover the full range of institute science. Another single panel, consisting of 12 members drawn mainly from the industry, carried out the evaluation of knowledge transfer (BBSRC, 1999a : 3).

The nine visiting groups per individual institutes, each of which comprised 12 members from the UK universities, business, government and where appropriate from overseas, considered the totality of the institute's research for the period April 1996 to December 2000 and the institute's forward plans for the period January 2001 to December 2004 (BBSRC, 1999a). In particular, the detailed task of visiting groups is (1) to carry out a scientific audit of the quality, (2) to conduct an institute's overall programmes and (3) to examine the institute's effectiveness in managing its resources to meet the needs of users and beneficiaries of its research (BBSRC, 2002a).

Finally, some members of staff in the Finance Group, one of the four departmental groups in the BBSRC,<sup>9</sup> examined the external income which each institute had received in a competitive way from April 1996 to March 2001 (BBSRC, 1999a).

## **(2) Evaluation Unit**

In the BBSRC, the evaluation unit, which was responsible for the organising the evaluation of institutes, was the "Assessment and Policy Unit" (APU) within the "Business Innovation and International Group". The Assessment and Policy Unit, in which was six members of staff as at January 2002, comprises missions such as (1) the liaison between government departments, (2) the assessment of institutes, (3) the management of research information and the Oasis database, (4) the preparation of output and performance indicators (OPIs) and (5) the evaluation of research programmes (BBSRC, 2002a).

Meanwhile, in the case of the 'cross-cutting review of science and research' and the 'quinquennial review of the six grant awarding research councils', the main role of evaluation unit, a temporary organisation for evaluation work which had been commissioned for the specific policy goals, was the mission of the secretariat. However, the Assessment and Policy Unit, one of the standing internal organisations of the BBSRC, carried out the actual evaluation of some evaluation objects as an internal evaluator as well as acted as the secretariat in the IAE 2001. Consequently, the overall co-ordination of the IAE 2001 was carried out by the Assessment and Policy Unit, with assistance from

---

<sup>9</sup> As at February 2002, the departmental group of the BBSRC comprised the following four groups : (1) the Business Innovation and International Group, (2) the Finance Group, (3) the Human and Corporate Resources Group and (4) the Science and Technology Group (BBSRC, 2002a).

the Science and Technology Group for the QS, the Finance Group for the CF and the Business and Innovation Unit for the KT.

### (3) Evaluation Information

For this evaluation, individual institutes to be evaluated were required to submit information, which was slightly different according to the each evaluation object, to the BBSRC Office. The detailed information required is summarised in <Table 4-11>.

<Table 4-11> Information Requirements in the IAE 2001

Evaluation Object		Information Requirements	Detailed Main Information
QS		Overall research staff summary	-
		Research leader details	-
		Published assessable outputs	- description & title of output - names of authors - name of journal & book publisher - volume number and pages
		Prizes and other esteems	- source of award, title, etc.
KT	Retrospective Assessment	Output & performance indicators (OPIs)	- studentships (PhD & Master) - user involvement - women in science - total publication output - external income - collaboration current - staff industrial consultancies - staff joining/leaving institute - exploitation of science (patents, plant breeder's right)
		Institute spin-off companies	-
		Written statements on the institute's activities in KT	-
	Prospective Assessment	Written statements on the future KT activities	- overall policy for promoting KT - policy for knowledge transfer - policy for developing an entrepreneurial culture - policy for meeting user needs - policy for promoting the public understanding of science etc.
	CF	Income & expenditures schedules	-
VG		The director's overview statement	- including organisational chart
		Individual research programme	- title, total cost, description, staff details and organisational chart, projects and funding, output, forward plan, etc.
		The institute's corporate plan	-

From this, it appears that even though the ratio of its usage was different according to the evaluation object, both quantitative information that could be more easily counted or calculated, and qualitative information that could be rather difficult to measure were required. Meanwhile, most of this information from the data already submitted was analysed by the BBSRC Office, through the existing research database known as "Oasis".<sup>10</sup>

As a result, it is evident that even though the information requirements differed depending on the evaluation object, both quantitative and qualitative information were adopted in the IAE 2001 exercise.

#### ***4.3.3.3. Evaluation Performance***

##### **(1) Evaluator**

As already noted, two types of evaluator were engaged in this evaluation. One was the BBSRC officers such as the staff of the APU and of the Finance Group, the other was external experts involved in panels and visiting groups respectively (BBSRC, 1999a).

Therefore, it is apparent that both internal and external evaluators participated in the IAE 2001. However, even though two kinds of evaluators were engaged, the external evaluators played more important roles in this evaluation, in a sense that they conducted the evaluation of the QC, the KT and the VG assessment which were regarded as the main evaluation objects.

##### **(2) Evaluation Method**

In the IAE 2001, qualitative methods such as the site visit and report reviews as well as quantitative methods including a bibliometrics were adopted as the main evaluation methods. While both evaluation methods were adopted, quantitative ones were more important than in other UK evaluation exercises such as the cross-cutting review and the quinquennial review. On conducting the retrospective assessment, when the performance of the quality of science (QS) and of the knowledge transfer (KT) was evaluated, a

---

<sup>10</sup> Oasis database contains information on all research programmes funded by the BBSRC at its sponsored institutes, universities and higher education institutions and other research institutes. This database covers details of research grants, postgraduate studentships, fellowships, research projects, the BBSRC-sponsored institute staff publications and so on (BBSRC, 2002a).

quantitative method including numbers of publications and of awards, and output and performance indicators (OPIs) was mainly used. In particular, in judging the quality of the published research articles, a bibliometric analysis was adopted as a measurement of scientific and technical excellence. Also, in order to evaluate the Competitive Funding, quantitative methods were mainly used (BBSRC, 1999a).

In contrast, in assessing the prospective aspect of the KT and the institute's performance as a whole, a qualitative evaluation method was adopted, in which the subjective judgement of the individual evaluator was more emphasised. In particular, when the visiting groups evaluated the overall performances of the institute, they spent 3 days at each institute. Each visiting group considered the added value provided by the institute, its distinctive contribution to the BBSRC's mission and the UK's science base, and the appropriateness and strategic relevance of the institute's forward plans. Reports from the visiting group were then considered by the Council (BBSRC, 2002d).

### **(3) Evaluation Criteria and Indicator**

Specific evaluation criteria and indicators, reflecting the attributes of each evaluation object, were adopted, as shown in <Table 4-12>.

In particular, in the evaluation of some objects, such as the quality of science (QS), competitive funding (CF) and knowledge transfer (KT) (especially retrospective assessment in the latter case), a quantitative measurement unit was used as the major evaluation indicator. On the contrary, in the case of the prospective assessment of KT and of the visiting group (VG) assessment that were rather difficult to evaluate in a quantitative way, qualitative criteria had to be utilised (BBSRC, 1999a).

Consequently, in this evaluation, it appears that, depending on the evaluation object, quantitative evaluation indicators were adopted as much as possible, where available.

**<Table 4-12> Evaluation Criteria and Indicators in the IAE 2001**

Evaluation Object		Criteria	Main Indicator
<b>QS</b>		<ul style="list-style-type: none"> <li>- the contribution of individual research leaders to their subject</li> <li>• quality of the publications or other research outputs</li> <li>• distinction of the research leaders</li> </ul>	<ul style="list-style-type: none"> <li>- number of publications                             <ul style="list-style-type: none"> <li>• ISI journal articles</li> <li>• refereed journal articles</li> <li>• conference proceeding/abstract</li> </ul> </li> <li>- number of awards</li> </ul>
<b>KT</b>	<b>Retrospective Assessment</b>	<ul style="list-style-type: none"> <li>- extent and appropriateness of exploitation of research</li> <li>- level of knowledge transfer</li> <li>- extent of institute involvement with users</li> <li>- level of staff involvement in exchanges and consultancies with industry</li> </ul>	<ul style="list-style-type: none"> <li>- studentship numbers</li> <li>- submission &amp; completion rate of studentships</li> <li>- governing body members from industry</li> <li>- collaboration cases with industry, university and other countries</li> <li>- number of staff exchanged with and involved in industry consultancies</li> </ul>
	<b>Prospective Assessment</b>	<ul style="list-style-type: none"> <li>- the adequacy of proposed arrangements</li> <li>- the influence of these arrangements on the future performance in KT</li> </ul>	<ul style="list-style-type: none"> <li>- evidence or measures for promoting and encouraging KT</li> </ul>
<b>CF</b>		<ul style="list-style-type: none"> <li>- the effectiveness of the BBSRC CSG to attract external income</li> </ul>	<ul style="list-style-type: none"> <li>- income from the BBSRC, the DEFRA, the EC and other funding sources</li> </ul>
<b>VG</b>		<ul style="list-style-type: none"> <li>- the value of the institute's overall contribution to the UK science base</li> <li>- the added value derived from the interdisciplinary and rolling nature of the research</li> <li>- the future direction of individual programmes of research</li> <li>- the institute's overall future strategy</li> <li>- the contribution to the BBSRC strategy</li> <li>- contribution of training</li> </ul>	-

#### **4.3.3.4. Evaluation Utilisation**

##### **(1) Evaluation Report**

The evaluation report of the IAE 2001 was produced by the Assessment and Policy Unit. This report covers some evaluation components such as an evaluation purpose, evaluation objects, evaluation methods, the research highlights of each institute and

evaluation utilisation. In particular, regarding the evaluation utilisation, it dealt with the detailed budget allocation of the CSG, capital grants and application limits for Responsive Mode and Initiative grants for each institute and HRI in the period of 2002-2004 (BBSRC, 2002d).

Although this report is expected to function as an important reference for the BBSRC-supported institutes as well as the research communities concerned in the future, it was not open to the public in the form of either electronics or of the hard copy. The summary of the outcome of this evaluation was merely posted on the BBSRC website (<http://www.bbsrc.ac.uk/about/centres/iae2001web.html>) (BBSRC, 2002d).

## **(2) Utilisation Type**

As in other evaluation work, the findings of the IAE 2001 were accepted as a policy means in relation to budget allocation as well as of the development in the management system and the research plan for both the institutes and the BBSRC itself in future. That is, the BBSRC Council used these results to set funding plans for the period of 2002-2006 and endorsed the key recommendations.

Firstly, on the basis of the results of the IAE 2001, the BBSRC Council agreed changes to the institute's Core Strategic Grants, adjusted the total values of the responsive mode awards for which the institute will be able to compete and made a number of specific capital and equipment grants. In particular, Council agreed the Core Strategic Grants for the BBSRC-sponsored institutes and for the Horticulture Research International (HRI) for April 2002 to March 2004 within the budget allocation established after the Spending Review (SR) 2000 exercise, involving an aggregate 4% increase per annum. Meanwhile, for beyond April 2004, only indicative planning figures for the BBSRC's core grant, namely average increase of 2.5% per annum, were provided. The reason was that there has been the uncertainty as a result of the government's review of agricultural policy and the likely associated shift of policy emphasis within the Department for Environment, Food and Rural Affairs (DEFRA)'s research programmes (BBSRC, 2002b ; 2002d ; 2002e). Their performance of the research expenditures will be monitored regularly by the Chief Executive and by the Council.

Secondly, it also may seem that the results of the IAE 2001 could be devoted to the



developments in the management system and to the research strategic plan of its sponsored institutes as well as of the BBSRC. Recommendations from all aspects of the IAE 2001 were provided to the institute directors, whose progress in implementing them is monitored annually by the Council. In addition, it is apparent that outcomes from the IAE 2001 will feed into the BBSRC's strategic planning and the research arrangements. In particular, they may influence the Council's ten-year scientific opportunities road map and the next five-year strategic plan. Hence, these will establish the BBSRC's ten-year vision of key opportunities in the non-medical biosciences and identify the Council's strategic priorities for the coming five years (BBSRC, 2002e).

In consequent, the findings of the IAE 2001 were fully utilised from the instrumental perspective and are also expected to be adopted as very essential policy means from the conceptual perspective.

#### ***4.3.4. Evaluation of S&T Programme : Evaluation of Funding for UK Civil Space Activity by the DTI and other Departments (2001)***

Between late 1999 and early 2001, the cross-departmental evaluation on UK civil space activities was conducted by the five funding partners, namely the DTI, the DETR (now DfT), the Meteorological Office (normally known as the Met. Office), the NERC and the PPARC.<sup>11</sup> This was the first evaluation of its kind in the UK space sector. The final report was published in July 2001. It comprised an overarching report prepared by Technopolis Group<sup>12</sup> and five separate evaluation reports provided by the funding partners.

---

<sup>11</sup> The funding partners' particular areas of interest are as follows : (1) The DTI is particularly concerned with providing support for the UK space industry to ensure that UK companies are in a strong position to compete in the space market. Its evaluation therefore focused on these industrial and commercial issues. (2) The DETR is both a funder of space activities and a user of data derived from those activities. The evaluation concentrated on analysing the programme costs and expected or actual benefits in both the monetary and qualitative perspectives. (3) The Met. Office is principally interested in the collection and use of space derived data to increase the accuracy of forecasts and potentially, the range of products it can provide. Its evaluation effort concentrated on trying to derive economic benefits for the Met. Office, customers and the nation. (4) The NERC is principally concerned with research on the national environment. Its evaluation concentrated on the extent to which developments in earth observation techniques and data processing and use have benefited environmental science. (5) The PPARC is particularly interested in enabling British universities to carry out top class, cost effective science. The evaluation concentrated on the scientific, industrial and public policy benefits of the PPARC's funding of space science (DTI/AU, 2001 : 205).

<sup>12</sup> Technopolis Group Ltd. is an international research and consulting company focusing on science, technology and innovation policy. The company is based in the UK, the Netherlands, France and Austria and has a staff of 30 plus a larger number of associate researchers. Its work spans most scientific disciplines,

Although some funding partners were involved together in this space sector evaluation, the Assessment Unit (AU) within the TESE (Technology, Economics, Statistics and Evaluation) Directorate of the DTI played the major role in this evaluation. For this reason, following the overview describing the background and the general information about this evaluation, emphasis will be placed on the analysis of the DTI's evaluation practices where appropriate, even though the overall evaluation work has already been described in broad terms.

#### **4.3.4.1. Overview**

The UK has a long history of pioneering involvement in space activity. Before the formation of the European Space Agency (ESA), the UK was a leading member of both the European Launcher Development Organisation (ELDO) and the European Space Research Organisation (ESRO). It has also continued to focus its funding for civil space through the ESA, being involved variously in space science, satellite telecommunications and earth observation (DTI, 2001c : ii). The UK government has focused its funding for space activities through the British National Space Centre (BNSC), a directorate of the DTI. The BNSC acts as the lead department with responsibility for co-ordinating all of the UK's civil space policy and activities. As the lead department, the BNSC has not only represented the interests of the DTI, but also those of a number of other government departments and research councils : the MoD, the DETR, the Met. Office, the PPARC and the NERC. Government spending on civil space activities amounts to approximately £190 million per year, the bulk of this being provided by the DTI and the PPARC. The remaining partners provide funding for individual space R&D programmes according to their own perceived needs and how the programmes contribute to them (Technopolis Group, 2001: 126).

Although the DTI has a long tradition of evaluating its industrial support programmes, there has never been a comprehensive examination of UK space programmes in its entirety. This evaluation was therefore the first comprehensive review undertaken of the impact and achievements of public support for UK civil space activities. It was a complex

---

technologies and industrial sectors (Technopolis, 2002). For more detailed information, see the Technopolis website (<http://www.technopolis.co.uk>). Technopolis was awarded the contract for this evaluation work following an open competition (DTI/AU, 2001 : 137) .

and wide-ranging review, covering a long period of time (some 25 years of ESA membership), and both international (principally European) and domestic activities (DTI, 2001c : ii). In late 1999, this evaluation was initially launched by the Evaluation and Policy Improvement Committee (EPIC) of the DTI. However, in order fully to capture the spectrum of benefits that derived from this funding, the evaluation was widened to encompass the full range of UK public funding for civil space activities. This was accompanied by each of the funding partners undertaking an evaluation together to give a comprehensive picture of the impacts of the UK public funding of space. So, the evaluation set out to consider – in equal measure – the industrial, scientific and public policy benefits that had been realised (Technopolis Group, 2001 : 11, 126).

This evaluation was carried out on the basis of the comprehensive self-assessments performed by the five funding partners,<sup>13</sup> using the common “evaluation protocol” prepared by the “evaluation team”. An evaluation protocol, which was translated from the policy questions by the DTI’s Assessment Unit, formed the basis of the five self-assessments and the subsequent terms of reference for Technopolis’ preparation of the overarching report. The evaluation protocol dealt with five classical evaluation questions, which were : (1) the appropriateness of the rationale, (2) the relationship between national programmes and the ESA, (3) goal attainment and impacts, (4) value for money in the outputs and inputs and (5) management efficiency (Technopolis Group, 2001 : 12). The evaluation team comprising specialist policy and evaluation consultants from both the Technopolis Group and three international space experts provided advice on the funding partners during the preparation of their reports and subsequently synthesised that material in an overarching report. This team also sought to ensure both that the evaluation questions had been answered and that the answers had been valid. For this evaluation, the BNSC Resources Board<sup>14</sup> also set up a Space Evaluation Steering

---

<sup>13</sup> The individual self-assessment reports were as follows (Technopolis Group, 2001: 10) :

- (1) DTI Financial Support for Civil Space Activity (DTI, October 2000, internal work)
- (2) Evaluation of UK Space Funding : Support for Global Navigation Satellite Systems (GNSS) and Advanced Along Track Scanning Radiometer (AATSR) (DETR, August 2000, internal work)
- (3) Evaluation of UK Space Funding (The Met. Office, August 2000, internal work)
- (4) NERC Evaluation of Civil Expenditure on Space (NERC, July 2000, commissioned work)
- (5) Evaluation of UK Civil Space Programme (PPARC, June 2000, internal work)

<sup>14</sup> The Resources Board supervises the BNSC activities as a whole. It comprises senior representatives from all of the funding partners and chaired by a senior DTI official, currently the Director General of the Business Competitiveness Group, who reports to the ministers responsible for the various funding partners. The funding partners are all represented on the Resources Board and most have staff seconded to the BNSC headquarters in London (Technopolis Group, 2001 : 18).

Committee (SESC) to oversee the evaluation process. The Steering Committee convened a Space Evaluation Working Group (SEWG) comprising the authors and analysts preparing the funding partner report (Technopolis Group, 2001 : 7, 10-11).

Meanwhile, the evaluation methodology adopted here was novel in as much as it combined self-assessments, which had been carried out by dominant specialists representing the constituencies under review, with a separate validation procedure involving R&D evaluation professionals and international space experts. The evaluation was performed in three phases. The first phase involved self-assessments carried out in parallel by the funding partners using the evaluation protocol to guide their work. In terms of data sources, the funding partners relied upon internal archives for the most part with a small number of comparative international data where these were available already. The second phase involved validation and elaboration of the funding partner reports through a process of mentoring provided by evaluation professionals, and structured assessment by international experts. The third phase involved an evaluation team in the preparation of a synthesis – the overarching report – which sought to ensure that the evaluation questions were answered as best they could be on the available evidence. The overarching report adopted a stringent position on the funding partners' statements, carrying over only that material which was corroborated by explicit evidence, the evaluation team's international comparisons or the international experts (Technopolis Group, 2001 : 12).

In order to ensure that the separate evaluations of the funding partners covered fundamental evaluation issues, an evaluation plan was developed and approved by the Working Group and Steering Group.<sup>15</sup> The plan set out the main issues <sup>16</sup> and suggested methodological approaches <sup>17</sup> that might be used to address them. It was sufficiently

---

<sup>15</sup> The space evaluation "Working Group" consisted of evaluators from the DTI and the funding partners plus representatives of the BNSC and OST. The "Steering Group" was chaired by the DG/OST and populated by the funding partners, the DTI TESE/Assessment Unit, the BNSC, the FRM and the UKISC (UK Industrial Space Committee) (DTI/AU, 2001 : 139).

<sup>16</sup> The main evaluation issues were as follows : (1) relationship between the ESA, the national and other (e.g. EUMETSAT) programmes, (2) the rationale, (3) market failure, (4) additionality, (5) counterfactuals, (6) benefits and beneficiaries, (7) objectives and (8) programme administration (monitoring and appraisal) (DTI/AU, 2001 : 207).

<sup>17</sup> The following list represented the methodological approaches which were expected to form the basis of the evaluation : (1) the mapping exercise on the space sector, (2) meetings with programme managers, (3) meetings with peers and experts within the industry, (4) structured interviews with direct beneficiaries and others, (5) review of existing files concerned and (6) review of national programme activities in other ESA

broad to allow each of the funding partners to identify the areas of particular importance to them whilst still covering the fundamental evaluation topics (DTI/AU, 2001 : 139).

As already noted, since one of the funding partners, the DTI was the initial organiser of this evaluation work and also bore the cost of it. In what follows, I shall confine the analysis of the case study to the DTI's evaluation exercise.

#### ***4.3.4.2. Evaluation Paradigm***

##### **(1) Evaluation Purpose**

The DTI evaluation was intended to examine the rationale for, and the impacts and achievements of, public funding of all UK civil space activities, including national and European Space Agency programmes. Thus, it concentrated on the industrial and commercial aspects of public funding of UK civil space activity (DTI/AU, 2001 : 138).

In other words, the evaluation sought to obtain answers to a number of key questions, principally concerning the appropriate level of funding for civil space activity and the balance between ESA, other international and national programmes, as set out in the agreed evaluation plan. These were as follows (DTI/AU, 2001 : 138) :

- What was/is the underlying rationale for government intervention in civil space activity? Is this still relevant?
- What degree of additionality has been/is achieved by public funding? What were/are the alternatives?
- What are the principal costs and benefits, and who are the principal beneficiaries, of public funding?
- Is the current level of public funding appropriate as between major application areas such as communications, earth observation, science and so on?
- Are the balance of funding between ESA/international programme and the national programme appropriate?

Additionally, policy makers needed to consider how future policy on space should be delivered and managed in order to maximise the potential gain from this activity for the

---

countries (DTI/AU, 2001 : 207).

UK, including how best to support and further small and medium sized enterprise (SME) involvement in the UK and international space supply chains (DTI/AU, 2001 : 139).

Consequently, this evaluation focused on the scale and types of impact which the public funding of the UK civil space activities has had. So, we may say that, in this evaluation, ensuring the accountability of the past expenditure was the most important evaluation purpose. In parallel, programme improvement in the space sector was also conceived as a main evaluation purpose.

## **(2) Evaluation Object**

The DTI evaluation was a complex and wide-ranging review, covering a long period of time since the mid-1980s (some 25 years of ESA membership) and both international and domestic activities (DTI/AU, 2001 : 137). Therefore, we may see that the level of the evaluation object normally covered the various programme-level works including pure science, applied research and development, the application of the research outputs and international collaboration in the space sector. When one more government departments supported a specific activity, it sometimes included policy-level work. Meanwhile, the scope of the evaluation object covered the home programmes mainly for the UK's own purposes as well as the international programmes, such as, in particular, collaborative R&D programmes between ESA members.

## **(3) Evaluation Type**

As already discussed, this evaluation was undertaken for two main purposes. One was to judge the impacts and achievements which had been derived from UK public funding for civil space activities. The other was to provide the programme improvements necessary for the future to the policy makers concerned. From these, it seems that the DTI work was carried out both as a retrospective evaluation, in the manner of an “output” and “impact” evaluation and also as a prospective “assessment”. Regarding the areas of outputs and impacts, the cross-departmental evaluation considered, in equal measure, the industrial, scientific and public policy benefits that have been realised, whereas the DTI evaluation emphasised, in particular, the industrial and commercial aspects of public funding (Technopolis Group, 2001 : 11 ; DTI/AU, 2001 : 138).

#### **4.3.4.3. Evaluation Resources**

##### **(1) Evaluation Manpower**

Two groups of manpower were involved in the DTI's evaluation. The first was the external evaluation team comprising eight members and the second was the internal evaluators consisted of three experts from the TESE. The evaluation team, which was awarded the contract for this work by open competition in April 2000, comprised five-specialist policy and evaluation consultants from the Technopolis Group and three international space experts (DTI, 2001c : 7). The external consultants of the Technopolis Group each had a background in one or more of pure science, engineering, humanities and a social science. They had also worked for the public sector, private companies including a consulting firm and international organisations such as the EC (Technopolis, 2002). The three international experts in the area of space or astronomy were invited from Sweden, Germany and the Netherlands. They were former members of related organisations, such as the European Space Agency, or current members of the European Southern Observatory. The main roles of the evaluation team were (1) to prepare an overall report covering all elements of government funding, (2) to draw conclusions and make recommendations for appropriate improvements, (3) to provide advice and guidance on aspects of evaluation methodology to the individual evaluators including the AU in the DTI, (4) to play an active part in meetings of the space evaluators' working group, (5) to maintain an awareness of the progress of individual evaluations through liaison with evaluators beyond the working group and (6) to recommend and draw upon other sources of information relevant to the overall evaluation (DTI, 2001c : 7 ; Technopolis Group, 2001 : 127).

Meanwhile, of the three evaluators in the DTI, while two came from the AU, one was from the Economists and Statisticians Unit, even though they were all staff of the TESE. They conducted the actual evaluation on the DTI's funding programmes for space under the guidance of the evaluation team and, sometimes, in collaboration with them (DTI/AU, 2001).

## (2) Evaluation Unit

This evaluation was carried out by the Assessment Unit (AU) which is part of the TESE.<sup>18</sup> As at February 2002, the AU contained five professional evaluators most of whom had a scientific or technological background. The present background of the AU evaluators includes chemistry, chemical engineering, information technology and aircraft engineering. Several of the evaluators also have experience in budget and programme management. In carrying out evaluation, the AU draws on the expertise of the four professional economists in the TESE and, in some cases, one of the senior economists may take the lead in undertaking the evaluation. Advice on surveys and statistical matters is available from one professional statistician (Barber, 1999a : 6 ; DTI, 2002a).

In addition to *ex post* evaluation, the AU also has responsibility for overseeing the monitoring of programmes within those budgets, such as Innovation Budget, Space Budget and Aerospace Budget, for which the TESE is responsible. As a result, the relative amounts of manpower which the AU devotes to programme appraisal, monitoring and evaluation, which sometimes includes members of staff from the 'Economists and Statisticians', can be seen from < Table 4-13> which shows projections for the financial year 1999-2000 (Barber, 1999a : 7-8).

<Table 4-13> Breakdown of the Manpower Commitment in the AU

Major Roles of the AU	Number of Persons a Year
Appraisal of new programmes	0.7
Continuous monitoring of programmes	1.2
Programme evaluation	3.1
Research associated with evaluation	0.8
<b>Total</b>	<b>5.8</b>
Of which	
Assessment Unit (AU)	5.0
Economists and Statisticians	0.8

<sup>18</sup> The TESE, based in London, is an interdisciplinary unit of the economists, statisticians, scientists and engineers within the Innovation Services (IS) Directorate of the DTI. It is made up of two groups who work closely together : (1) the Assessment Unit who are responsible within the DTI for *ex post* evaluation of innovation and technology support programmes, provide advice on the formulation of new programmes and supervise programme monitoring. (2) Economists and Statisticians who provide professional advice on science, technology and innovation and certain high technology sectors within the DTI, including the OST, provide a professional input to the evaluation work of the AU. As at February 2002, the TESE comprised 12 members of staff (DTI, 2002a).



### **(3) Evaluation Information**

The following sources were utilised to provide the information with which to cover the key evaluation questions outlined above (DTI/AU, 2001 : 140) :

- a newly commissioned and comprehensive study of the nature and extent of the UK space sector, its supply chain and key trends
- meetings with the BNSC programme managers and other experts within the funding organisations and, where appropriate, outside
- meetings with peers and experts within the industry
- an extensive structured interviews with direct funding beneficiaries such as recipients of government and ESA funding, indirect beneficiaries and end users
- a review of archived correspondence and files
- a review of the way in which other ESA member states provide support to their space industry, for comparison with the UK approach

From these sources, we may see that although some quantitative information was utilised in the actual process of evaluation, where appropriate, the evaluator of the AU took qualitative information as the main source for evaluation.

#### ***4.3.4.4. Evaluation Performance***

##### **(1) Evaluator**

Even though the external consultants such as Technopolis and Scientific Generics were fully involved in the funding partners' evaluation through preparing the overarching report and conducting a mapping study of the UK space respectively, the AU and its three evaluators took responsibility for the evaluation of the space programmes funded by the DTI (DTI/AU, 2001). They produced the overall evaluation protocol for the space evaluation, carried out the actual evaluation on the DTI funded activities and prepared its own separate evaluation report. Therefore, it is apparent that in fact the internal evaluator undertook this evaluation.

##### **(2) Evaluation Method**

In this evaluation, the DTI AU adopted diverse evaluation methods, on the basis of the common evaluation protocol which had been prepared by the evaluation team. Each evaluation method explained below was devoted to the achievement of one evaluation

purpose or, sometimes, of multiple evaluation purposes (DTI/AU, 2001 : 214-216).

(1) Mapping exercise on the space sector : The BNSC has already conducted a rolling programme of market studies, which provide useful information about this subject. However, further supplementary work by an external consultant was required. To investigate and describe the nature, extent and characteristics of the UK space sector, an industry mapping study was undertaken by consultants from Scientific Generics.<sup>19</sup> The scope of the exercise took account of the BNSC's studies in this area, the most recent of which had been undertaken in 1998. The results of this study by Scientific Generics provided a greater understanding of the UK space industry and assisted the evaluation team in identifying technologies, organisations and specialists who made a contribution to the evaluation. In particular, this mapping exercise made an attempt to identify and describe the following matters :

- the size and features of the overall UK space sector
- how the sector has changed over the last 10 or 15 years
- the work currently being done in relation to space technology
- the sort of companies involved
- what specific companies are doing
- which other companies could be involved
- information on the supply chain
- the benefits of technology transfer into non space industries

(2) Meetings with the BNSC's programme managers : The evaluator had a lot of meetings with appropriate programme managers and officials in the BNSC. Through these meetings, both the evaluation team and the AU were able to provide the programme managers with the approach of the evaluation and obtain some necessary information and project documentation. Such meetings were also important in ensuring that the evaluator fully understood the issues within the programme and the relationships between the various funding bodies. Meanwhile, programme managers were kept informed of the progress of the evaluation.

---

<sup>19</sup> Scientific Generics, a part of the Generics Group, which is based in the UK and has branches in the USA, Sweden and Switzerland, is a consulting company in the area of engineering, materials, life science, telecommunications, business innovation and electronics (Generics, 2002).

(3) Meetings with peers and experts within the industry sector : In order to obtain as unbiased a view as possible of the immediate and longer term impacts of the programmes, the evaluator contacted a number of experts within the space sector who had not been directly involved in any of the projects. Evaluators intended to co-ordinate such visits in order to minimise the burden on contributors.

(4) Structured interviews with direct beneficiaries such as recipients of government or ESA funding, indirect beneficiaries and end users : The AU also conducted a series of extensive interviews with representatives of industrialists, academics and customers/end users, in order to get their views on and experience of publicly funded space programmes, as well as to obtain feedback on the wider issues of commercial exploitation, subsequent developments, and wider industrial benefits. The interviews were conducted over a nine month period from November 1999 to July 2000 and embraced the full range of organisations delimited by size (employment and turnover) and space market segment. In order to obtain maximum benefit from the interviews, with a minimum burden on the contributors, a co-ordinated approach was developed with the other evaluators, including joint visits and the development of an agreed questionnaire covering the core evaluation issues. Interviewees included both direct and indirect beneficiaries and those who might be expected to have an informed view of the UK's space programme. For these structured interviews, the evaluation team contacted a total of 37 organisations and 58 individual interviewees.

(5) Review of files : A file research was utilised as a source of information, in particular, given the time period which the evaluation intended to cover, to get a historical perspective as well as to supplement information from other sources.

(6) Review of space programme activities in other ESA countries : The nature, scale and rationale for space funding in ESA member countries were investigated to provide a comparison for UK space funding. Comparisons were made in terms of the ways in which support has been provided and for what purposes and whether different methods of support have produced greater or more rapid benefits from the investment concerned. The evaluation used as a comparison, other ESA members that might be regarded as being of a similar profile to the UK in terms of size of contribution and participation in ESA programmes.

Therefore, it seems that, in this evaluation, a field survey, a literature review and an in-depth interview, which are regarded as the typical types of qualitative methods, were adopted as the key evaluation methods.

### **(3) Evaluation Criteria**

Space activity has normally many characteristics of a public good. So, without public sector financial support, the level of investment and resource employment in the activity are likely to be less than socially optimal. However, it is not possible to measure by how much. Therefore, many of the benefits claimed from continued space activity are difficult to trace, quantify or value (DTI/AU, 2001 : 134).

For this reason, in the DTI evaluation, rather qualitative evaluation criteria were adopted to judge the benefits and impacts of public funding. The main evaluation criteria, which could be drawn from the key policy questions noted above, are as follows : (1) the “rationale” of the government intervention in civil space activity and its current “relevance”, (2) the “additionality” of achievements of public funding and the existence of “alternatives”, (3) the principal “beneficiaries” of public funding, (4) the “appropriateness” of public funding among the supporting programmes, such as communications, earth observation, science and so on and (5) the “balance” between international programmes and national programmes (DTI/AU, 2001 ; 139) :

#### ***4.3.4.5. Evaluation Utilisation***

##### **(1) Evaluation Report**

The overall evaluation report was the first one to be published in full on the TESE website (<http://www.dti.gov.uk/tese>). It can also be found on BNSC’s website (<http://www.bnsc.gov.uk>) together with their responses, and has also been available free on CD-ROM and in hard copy. The report contained six chapters. The opening chapter was an overarching report that summarised and analysed the five subsequent separate evaluations of space funding by BNSC’s funding partners (DTI, 2002a). The remaining chapters included the separate evaluation reports which had been undertaken by the internal evaluation unit or the external evaluator under contract, as in the case of the NERC (DTI, 2001c).

Meanwhile, the AU also published its own separate evaluation report which comprised

one chapter of the overall space evaluation report. It was made up of 10 sections and three annexes.<sup>20</sup> In particular, in the section on “the impacts and achievements” of the DTI’s financial support, the evaluator in the AU grouped the principal industrial and commercial benefits of interest to the DTI into the following five categories : (1) the level of employment, (2) the quality of employment, (3) the turnover including ESA returns, (4) the technology and knowledge transfers and (5) the start-up companies (DTI/AU, 2001 : 142-145). Then, in order to discuss and suggest the supporting evidence for each issue, the AU utilised quantitative as well as qualitative information.

## **(2) Utilisation Type**

In both the overarching report and the AU report, there were a lot of recommendations which had been initially drawn from each funding partner’s evaluation report. Most of the recommendations were approved by the BNSC Resources Board.

In the overarching report, Technopolis acted as an evaluation team which synthesised the findings of the separate evaluations by the funding partners and made various recommendations as follows :

- Upgrading the UK’s strategic management systems
- Revisiting the merits of creating a budget holding executive agency
- Reviewing the funding arrangements
- Preparing ROAME statements and implementation plans
- Improving the evidence base, including preparing the official statistics
- Exploring ways in which to improve exploitation of space investments
- Maintaining the current emphasis on programme delivery through the ESA
- Reviewing the ESA’s financial strategy and budgetary rules
- The UK should continue to foster EC-ESA co-operation

All of these issues, except one, were nearly accepted by the government via the BNSC

---

<sup>20</sup> The detailed title of each section is as follows (DTI/AU, 2001) :

- (1) Introduction
- (2) Evaluation approach and method
- (3) The impacts and achievements of the DTI financial support for UK civil space activities
- (4) UK funding for civil space activity
- (5) Rationale and principal aims for government funding of civil space activity
- (6) The European space industry
- (7) The UK space industry
- (8) Benefits for companies
- (9) Programme management
- (10) Conclusions and recommendations

Resources Board. The suggestion rejected was that BNSC partners should replace their own programme approval and management process arbitrarily with the DTI's ROAME system (BNSC, 2002 : 1-6).

Meanwhile, the AU also made a lot of recommendations, which sometimes slightly overlapped those of the evaluation team. Certain individual recommendations were fully accepted by the government through the DTI. These were as follows (BNSC, 2002 : 7-8) :

- The rationale for a measure of DTI support for civil space activity and for continued membership of the ESA is supported by sound economic arguments. The BNSC should use these economic arguments to enhance the case for public funding.
- To capitalise fully on the scale and scope advantages offered by continued ESA membership, a real increase in funding for ESA programmes might be worthwhile. The DTI should give full consideration to the case for increasing spends within the ESA.
- A strategic review of the commercial opportunities for civil space activity should examine the case for a more balanced portfolio of complementary space investments within the UK budget to avoid reliance on a single technological force. This should be used to inform the next BNSC review of UK space policy.
- The BNSC should make greater efforts to ensure that the purpose of the DTI domestic space budget is widely understood and accessible throughout the UK space industry.
- To conduct an independent expert and in-depth study, within an appropriate timescale, of the process, scale and value of the transfer of technologies (product, process and knowledge) from the UK space sector to other industrial sectors.
- The BNSC should continue to seek the views of all interested parties and outside experts to ensure that UK space policy is properly focused, clearly understood by all and has strategic vision.
- The BNSC must set in place improved financial and programme-monitoring system in line with the continuous monitoring requirements expected of other DTI directorates.
- There should be regular (3-5 yearly) audits and evaluation of space programme spending and outcomes.

In consequence, it is apparent that most of the findings of the civil space evaluation were accepted by the related government departments. As a result, they will be expected to be

directly and properly reflected in the policies and strategies of the government departments which intend to enhance the space activities under their responsibility in the future.

#### ***4.3.5. Evaluation of R&D Programme : Review of the Gastrointestinal Response to Food (GRF) Programme by the BBSRC (2000)***

The BBSRC supports a number of research areas through specialised initiatives or co-ordinated programmes where the need for increased effort for scientific and strategic reasons has been identified. The BBSRC's research programmes underpin industries including the agricultural, bioprocessing, chemical, food, healthcare and pharmaceutical sectors. It is the BBSRC's policy to evaluate all of its research programmes and to use the conclusions to inform future funding decisions. According to this policy, in May 2000, the BBSRC evaluated the achievements of the "Gastrointestinal Response to Food (GRF)" programme which had been started in October 1992.

##### ***4.3.5.1. Overview of the Programme Management in the BBSRC***

The BBSRC funds research that is aimed at increasing human understanding of how living organisms function and behave, from the level of individual molecules and cells to populations and ecological systems, and at providing knowledge that can be used to develop new technologies and products for medicine and industry, and for better management of natural resources. The research encompasses fundamental, strategic and applied areas of study and covers a broad range of biological and related disciplines, for example, from biological chemistry, the biology of bacteria and viruses and computer modelling of proteins and other large molecules to studies on nutrient cycling, animal welfare and evolution. For organisational purposes, the research programmes of the BBSRC cover seven key areas which are administered by seven standing research committees <sup>21</sup> : (1) agri-food, (2) animal sciences, (3) biochemistry and cell biology, (4)

---

<sup>21</sup> The seven standing research committees have been set up to advise the Strategic Board. The scientific remits of the committees are designed to cover the full range of science supported by the BBSRC, including both basic and strategic science. Each committee consists of twelve members drawn from the academic and user communities. The committees are responsible for (1) peer review of research proposals, (2) committee studentship allocations, (3) programme planning, including identifications of low priority areas in which funding might be reduced and (4) project evaluation (BBSRC, 2002a).

biomolecular sciences (joint with the EPSRC), (5) engineering and biological systems, (6) genes and developmental biology and (7) plant and microbial sciences (BBSRC, 2002c).

The BBSRC funds research through a range of different mechanisms, one of which is the provision of research grants for specified projects. These research grants are categorised as follows : (1) Responsive Research Grants : These are awarded in response to unsolicited research proposals in any area relevant to the BBSRC's mission. As a general principle, the BBSRC allocates as much research grant funding as possible in responsive mode, assessed by research committees. These grants are open to UK universities, the BBSRC-approved academic analogues and the BBSRC-sponsored institutes. (2) Responsive Research Initiatives : These involve a call for proposals with applications invited with one or more closing dates. They are normally used in response to an emerging scientific opportunity where there is a need to enhance research activity to stimulate a particular area of science within the established BBSRC science base. (3) Thematic Research : The BBSRC decides whether, for any of its proposed Responsive Research Initiatives, there are grounds for extending access beyond UK universities, the BBSRC-approved academic analogues and the BBSRC-sponsored institutes and so bring to the scientific area added strength or expertise in order to deliver the research required. Where this is decided, the initiative is labelled "Thematic" and open to applications from institutes and units of the other research councils, the Scottish Agricultural and Biological Research Institutes (SABRIs), the government research establishments, government agencies and UK non-profit research organisations (BBSRC, 2001c : 2).

Each committee is responsible for the development of its own programme of funded science, and for developing those areas of science that it considers should have priority. The Agri-Food Committee supports basic and strategic research on the quality and safety of the food materials through the food supply chain, as well as the sustainability and efficiency of processes involved in the primary production, manufacture, storage and distribution of food and food ingredients. Its research portfolio comprises some broad themes and a number of specific priorities (BBSRC, 2001c).

The Gastrointestinal Response to Food (GRF) programme has been one of the research priorities in the Agri-Food Committee. The GRF programme has addressed the molecular



biology, biochemistry and physiology of nutrient interaction with the gut in relation to human health and nutrition, including systematic aspects. The area includes research into the regulation of nutrient uptake and cell signalling in relation to intracellular processes, intercellular interactions between cell populations and bacterial cells, enteric/vagal/endocrine control of the gastro-intestinal tract and the role of the gut flora (BBSRC, 2002c). As at May 2000, a total of 31 grants were awarded to 26 principal investigators (PIs) under the GRF programme.<sup>22</sup> The total funding for these awards was £6,434,763 (BBSRC, 2000 : 13).

#### ***4.3.5.2. Evaluation Paradigm***

##### **(1) Evaluation Purpose**

In the BBSRC, all research programmes need to be evaluated to assess (1) achievements of their research aims and objectives, (2) contributions to the subject area as a result of the programme, (3) benefits from the programme being co-ordinated centrally by the BBSRC, (4) added value in awarding grants through the programme and (5) the level and quality of the outputs from the programme (BBSRC, 2001b).

In the case of the evaluation of the GRF programme, the evaluation purpose indicated above should be the terms of reference. So, it seems that the evaluation of the GRF programme was also concentrated on the “accountability” and the “programme improvement”, similar to the evaluation of other R&D programmes.

##### **(2) Evaluation Object**

As already indicated, as at May 2000, the GRF programme comprised 31 grants covering research carried out by the 26 principal investigators (BBSRC, 2000 : 1, 13). This means that, at the stage of its implementation, this programme was grouped, where appropriate, into specific R&D projects which were more relevant. Although, in principle, this evaluation was intended to examine the programme itself as a whole, the actual evaluation work needed to be performed on the basis of the direct outcomes of individual R&D projects. Therefore, it appears that the level of the evaluation object was the individual projects which had been consisted of the GRF programme.

---

<sup>22</sup> As at May 2000, additional grants were also awarded, but these have not yet started. So, they have not been included in this evaluation (BBSRC, 2000 : 13).

Meanwhile, as at January 2002, research work covered by seven grants was still being performed in this priority area, because this programme has been funded since 1992 (BBSRC, 2000 : 13). As the evaluation was conducted in May 2000, two projects, which had not yet started, were excluded from it. Thus, the scope of the evaluation object was confined to projects which had already been completed before May 2000 or had been running in the period between October 1992 and May 2000.

### **(3) Evaluation Type**

This evaluation was mainly aimed at examining and judging what the programme and actual specific individual projects for which grants had been awarded, had achieved in terms of programme management, collaboration work, research output and so on (BBSRC, 2000). It is apparent that this evaluation was conducted as an “output” evaluation and, where appropriate, as an “impact” evaluation, which were regarded as the typical evaluation types in the evaluation of the R&D sector.

#### ***4.3.5.3. Evaluation Resources***

##### **(1) Evaluation Manpower**

This evaluation was performed by two groups of evaluation manpower : “members of staff in the Assessment and Policy Unit” (APU) and “the review panel”. Members of staff in the APU issued a questionnaire to all grantholders and analysed their responses. They also conducted the bibliometric analysis of the research outputs. Then, the results of the analysis on the questionnaire and the bibliometrics were submitted to the review panel for further consideration and comprehensive synthesis (BBSRC, 2001b).

In general, a review panel is organised by the research committee which co-ordinates the research programmes. The panel consists of several members of the research committee plus other independent members, with the BBSRC staff as observers. The role of the review panel is to consider the questionnaire responses and bibliometric analysis as well as to review the scores of individual grants from the final report and, sometimes, including the progress report (BBSRC, 2001c). For the evaluation of the GRF programme, in practice, a total of 12 external experts, including some members of the Agri-Food Committee, joined the review panel.

## (2) Evaluation Unit

Both the “Assessment and Policy Unit” within the Business Innovation and International Group and the “Agri-Food Branch” within the Science and Technology Group <sup>23</sup> were involved in this evaluation (BBSRC, 2002a).

As already described in the case of the IAE 2001, in the BBSRC, the “Assessment and Policy Unit” has the responsibility for carrying out the evaluation of the research institutes and research programmes. Regarding this evaluation, it conducted the actual analysis of the questionnaires returned from grantholders and the bibliometrics of the research outputs. As at January 2002, it comprised six members of staff (BBSRC, 2002a), although the evaluation was carried out by one member of staff only.

By organising the review panel, the “Agri-Food Branch”, which is in charge of the co-ordination of the Agri-Food Research Committee, played a direct part in the evaluation. The “Agri-Food Research Committee”, one of the seven research committees in the BBSRC, is responsible for the final result of the evaluation, by considering the conclusions of the review panel. The members of the committee are normally drawn from academia, industry, government and research institutes, in order properly to represent the interests of the diverse scientific communities. They normally meet three times a year. A breakdown of the committee by membership type as at January 2002 is given in <Table 4-14> (BBSRC, 2002a).

<Table 4-14> Membership of the Agri-Food Committee in the BBSRC

<b>Organi. Role</b>	<b>Academia</b>	<b>Industry</b>	<b>Research Institute</b>	<b>Government Department</b>	<b>Total</b>
<b>Chair</b>	1	-	-	-	1
<b>Member</b>	5	5	1	-	11
<b>Observer</b>	-	-	-	3 *	3
<b>Total</b>	6	5	1	3	15

\* These are the DEFRA, the Food Standards Agency and the Scottish Executive Environment and Rural Affairs Department (SEERAD).

<sup>23</sup> Within the Science and Technology Group, there are five branches, which are responsible for the coordination of seven research committees : (1) the Agri-Food Branch, (2) the Animal Sciences Branch, (3) the Genetics and Biochemistry Branch which incorporated the ‘Biochemistry and Cell Biology Committee’ and the ‘Genes and Developmental Biology Committee’, (4) the Biomolecular Sciences Branch and (5) the Plant and Engineering Branch which incorporated the ‘Plant and Microbial Sciences Committee’ and the ‘Engineering and Biological Systems Committee’ (BBSRC, 2002a).

Meanwhile, the “Agri-Food Branch” acted as the secretariat to the committee, co-ordinating its work. As at January 2002, there were eight members of staff : one head, three programme managers, one administrative manager and three administrators (BBSRC, 2002a). As a result, the actual role of the committee and the branch in this evaluation was to arrange and support the review panel, which had taken into account both the questionnaire analysis and the bibliometric analysis conducted by the APU in its deliberation.

### **(3) Evaluation Information**

In this evaluation, three types of sources were utilised to gain the information : (1) a questionnaire returned from the grantholders, (2) a bibliometric analysis of the research outputs and (3) final grant reports form. In particular, the information obtained from questionnaires was amalgamated with details of outputs from the final report forms.

The questionnaires were distributed to the 31 grantholders to gain information relating to issues such as whether the programme had met its objectives, how well it had been managed by the BBSRC, details of new contacts and collaborations, further funding applied for and received and so on (BBSRC, 2001b). The detailed information that it was intended to obtain from the questionnaire is indicated in <Table 4-15>.

In order to get the information about both the quality and level of outputs from the GRF programme, a bibliometric analysis was conducted. In practice, in the bibliometric analysis, the emphasis was given to papers in the peer-review journals. These were published in 41 different journals, including the Biochemical Journal, the British Journal of Nutrition, the Clinical Science, Gastroenterology, the Journal of Lipid Research, the Journal of Physiology and Pharmacology, Molecular Pathology and Neuroscience (BBSRC, 2000 : 13)

The final report and interim report were also important sources for comprehensive information at both a qualitative and a quantitative nature (BBSRC, 2001b). A final scientific report form covering publications, new information of gene sequences and so on, must always be submitted within three months of the completion of the grant period. In some circumstances, that is, in the case of larger or longer-term grants and responsive research initiatives or thematic research programmes where there are special conditions

associated with research grants, the interim reports or site visits may be required.

<Table 4-15> Detailed Information from the Questionnaire for GRF Evaluation

Category	Information Intended to Gain
<b>General</b>	1. Do you have any comments on the organisation of the programme as a whole? 2. Do you think the programme achieved its aims in fostering links between scientists and clinicians in the area of gastroenterology? 3. Do you consider that the programme has enhanced the progress of your research?
<b>Meetings</b>	4. Did you attend the programme meetings/workshops? If yes, please list the meeting dates and locations of those you attend 5. Were the workshops/meetings (a) useful, (b) interesting but not useful, (c) a waste of time? If (a), in what way did you find it useful? If (b) or (c), please suggest how it might have been made more useful? 6. Why were you unable to attend the workshops/meetings?
<b>Networking/Linkage</b>	7. Has participation in the programme led to new contacts (both academic and industrial) being made : (a) within the UK? (b) overseas? If yes, please give details of the nature of the contacts. 8. Have you developed any existing contacts (both academic and industrial) further as a result of participating in the programme? If yes, please give details. 9. Have you established any new collaborations (both academic and industrial) as a result of participating in the programme? If yes, please give details of the nature and extend of the collaborations.
<b>Funding</b>	10. (A) Have you sought funds from the BBSRC or elsewhere to continue or develop the work funded by the GRF programme? If no, please give reasons. (B) Are you planning to seek funds to continue or develop the work funded by this programme? If yes, please provide details. 11. (A) If you have been successful in securing funding to continue the research carried out under your GRF grant, please provide details. If possible, please indicate the funding source and approximate value of the awards. (B) Please give details of the BBSRC grant applications that did not result in funding. 12. Has the GRF grant allowed you to extend your research activities into other areas? If yes, please give details.
<b>Outputs</b>	13. On a separate sheet, please provide an update of the information on publications given the end of award final report. 14. On a separate sheet, please list any other significant outputs from the research.
<b>Other</b>	15. Please feel free to express your views on any other aspects of the programme.

Source : BBSRC (2000), pp.2-6.

Generally, final reports are assessed by two suitable experts who consider (1) the success of the research in meeting its original objectives, including any agreed amendments, (2) the quality of the work carried out under the research grant, (3) the extent to which it has advanced knowledge and/or enhanced UK industrial competitiveness or quality of life

and (4) the general scientific and practical significance and specific outputs, such as publications, patents, collaborations with other scientists in industry or academia (BBSRC, 2001c : 15-16).

Consequently, in the evaluation of the GRF programme, while qualitative and quantitative information were equally utilised, quantitative information was considered more important than in other UK evaluation cases.

#### ***4.3.5.4. Evaluation Performance***

##### **(1) Evaluator**

While the internal evaluators (the APU staff) and the external evaluators (the review panel) were both involved in this evaluation, the internal evaluators played a more important role than the external ones. The staff of the Assessment and Policy Unit conducted a questionnaire analysis and a bibliometric analysis which are regarded as the main evaluation methods in the area of R&D project evaluation, whereas the review panel considered and reviewed the results from the Assessment and Policy Unit's work (BBSRC, 2001b). In practice, it seems that the mission of the review panel was concentrated on the synthesis of the draft of the evaluation outcomes, rather the actual execution of evaluation.

##### **(2) Evaluation Method**

Regarding the evaluation methods, both the analysis of the grantholders' view obtained from the questionnaire and the bibliometric analysis of the research outputs were undertaken. The questionnaire was distributed to the 26 principal investigators (PIs) in February 2000 with the aim of gaining participants' views on a range of issues, including (1) the co-ordination of the programme, (2) the usefulness of the workshops, (3) the usefulness of the programme for networking and (4) attracting additional funding. Grantholders were also requested to provide an update of their research outputs arising as a direct result of their GRF grant. The Assessment and Policy Unit conducted the actual analysis of the questionnaire on the basis of the views of 17 respondents out of a total of 26 PIs (BBSRC, 2000 : 1).

A bibliometric analysis of the research outputs was also conducted for this evaluation.

Calculations were based on research outputs given on the final and interim report forms in addition to the supplementary information extracted from the questionnaires returned by 17 of the PIs. The research outputs considered included (1) papers in peer-reviewed journals, (2) other significant publications and presentations such as non-refereed papers, review articles, book chapters, conference proceedings/abstracts and websites, (3) further grants obtained, (4) public understanding of science activities and (5) exploitation potential.

In particular, while the calculations of the papers included in both peer-reviewed and non-refereed journals, a greater weight was given to papers published than to those either in press or merely submitted. As a result, just as the GRF programme generated a total of 108 refereed publications including 44 papers either in press or merely submitted, so the ratio of the refereed publications to research grants was 2.1. In addition to these, another 181 outputs, such as non-refereed papers, book chapters, conference proceedings and so on arose from these grants (BBSRC, 2000c : 13-14).

### **(3) Evaluation Criteria and Indicator**

Even though there were no clear descriptions, it was possible to extract the evaluation criteria and indicators from the many official documents in the BBSRC. First of all, it was apparent that the results of questionnaire analysis and bibliometric analysis could be used as the main sources (BBSRC, 2000). Next, in general, as already indicated in the “evaluation information”, it became clear that the assessment criteria of the final report of the grant could be utilised as an important reference for the evaluation criteria (BBSRC, 2001c : 15-16). In addition to these, we were also able to extract references from the appraisal criteria indirectly. In the BBSRC, scientific excellence and cost effectiveness, which does not mean cheapness, are the essential criteria in the consideration of all proposals and, within an appropriately balanced portfolio, priority will be given to proposals which are also strategically relevant, underpin improvement in the quality of life and/or may contribute to creation of prosperity (BBSRC, 1999b : 25).

The evaluation criteria and indicators ascertained through the above means and actually adopted in this evaluation are summarised in <Table 4-16>.

**<Table 4-16> Evaluation Criteria and Indicator for GRF Programme**

<b>Evaluation Category</b>	<b>Evaluation Criteria</b>	<b>Evaluation Indicator</b>
<b>General</b>	<ul style="list-style-type: none"> <li>- the achievement of aims and objectives</li> <li>- the enhancement of research progress</li> <li>- the extent of advanced knowledge</li> <li>- cost effectiveness</li> </ul>	-
<b>Contribution</b>	<ul style="list-style-type: none"> <li>- UK industrial competitiveness</li> <li>- the quality of life</li> </ul>	-
<b>Securing Funding</b>	<ul style="list-style-type: none"> <li>- the generation of new funding through the programme</li> </ul>	<ul style="list-style-type: none"> <li>- the success of securing new funds</li> <li>- the size and nature of funds from the BBSRC and elsewhere</li> </ul>
<b>Collaboration</b>	<ul style="list-style-type: none"> <li>- the usefulness of workshops/ meetings</li> <li>- the usefulness of networking/ linkage</li> </ul>	<ul style="list-style-type: none"> <li>- attendance and its frequency of workshops/meetings</li> <li>- the existence of new contracts and collaborations with academia and industry</li> </ul>
<b>Research Outputs</b>	<ul style="list-style-type: none"> <li>- a scientific excellence</li> <li>- a practical significance</li> </ul>	<ul style="list-style-type: none"> <li>- the total number of published paper</li> <li>- the number of papers in the peer-review journals</li> <li>- the ratio of refereed publication to research grant</li> <li>- the number of PIs taking part in the activities for the public understanding of science</li> <li>- the number of PIs considering the potential for exploiting the results of their research</li> </ul>

#### **4.3.5.5. Evaluation Utilisation**

##### **(1) Evaluation Report**

The evaluation report was prepared by the Assessment and Policy Unit. This report contains the general information including the programme structure, the questionnaire sheets, the list of peer-reviewed articles and other research outputs, as well as the findings drawn from the questionnaire analysis and the bibliometric analysis (BBSRC, 2000).

This report was circulated to BBSRC staff and members of the review panel including the research committee members concerned. However, it was not made available to the



public, either electronically on the BBSRC website or in hard copy form.

## **(2) Utilisation Type**

In general, the findings of evaluation have been utilised for decision making at the level of existing individual R&D projects to determine whether they should be continued, reduced or terminated (BBSRC, 2002a). In line with the general policy of the BBSRC's Council, evaluation results of the GRF programme have also been used to identify priority areas for future funding (BBSRC, 2001b).

### ***4.3.6. Synthesis and Conclusion***

In this part, case studies about five evaluation exercises in the UK S&T and R&D sectors have been conducted. The synthesis of these case studies is given in <Table 4-17>, on the basis of the analyses of evaluation exercises. As a result, it seems that, generally in the UK, evaluation exercises in these sectors have similar characteristics, even though there are slight difference in some aspects of meta-evaluation components such as evaluation resources and evaluation performance, according to the evaluation purposes. While external evaluators played major roles in the 2000 cross-cutting review of science and research as well as in the 2001 quinquennial review of the six research councils, internal evaluators had more important missions, even though they were in collaboration with external evaluators, in the evaluation of both the UK civil space activities by the DTI and the GRF programme by the BBSRC. Also at the higher level of evaluation exercises such as the 2000 cross-cutting review of science and research and the 2001 quinquennial review of six grant awarding research councils, more qualitative evaluation methods were adopted. However, in the case of evaluation exercises at the project-level such as the GRF programme, the emphasis was placed on rather quantitative evaluation methods. In particular, it is apparent that, in all evaluation exercises, the findings of evaluation have been fully reflected in decision-making processes in relation to future budget allocation and further grants.

<Table 4-17> Synthesis of the Evaluation Exercises in S&T and R&D Sector of the UK

Components \ Cases		Cross-cutting Review (2000)	Quinquennial Review (2001)	IAE (2001)	UK Civil Space Activity (2001)	GRF Programme (2000)
Evaluation Paradigm	Purpose	- accountability - policy/programme improvement	- accountability - improvement of management in RCs - knowledge enhancement	- accountability - improvement of fund expenditures	- accountability - programme improvement	- accountability - programme improvement
	Object	- SET policies - related programmes - science and research in England	- past performance of RC - future management plan - knowledge enhancement	- QS, KT, CF - institute as a whole (VG)	- programme-level - policy-level, available - both domestic and international activity	- programme itself (in principal) - individual projects (in practice)
	Type	- output evaluation - impact evaluation	- output evaluation - assessment for future plan	- output evaluation - assessment for research plan	- output evaluation - impact evaluation - prospective assessment	- output evaluation - impact evaluation
Evaluation Resources	Manpower	- Enterprise Team (1) - steering group (5)	- steering group (17) - review team (6) - working group (85)	- 7 expert panels (84) - 9 visiting groups (108) - BBSRC's officers (6) (APU staff etc.)	- evaluation team (8) (Technopolis and inter. space experts) - TESE in the DTI (3)	- Assessment & Policy Unit (6) - review panel (12) - Agri-Food Branch (8)
	Unit	- Enterprise Team in HM treasury	- temporary mini-Direct. under SEBG in OST	- Assess. & Policy Unit	- Assessment Unit within TESE	- Assess. & Policy Unit - Agri-Food Branch
	Information Sources	- existing literature - consultation	- extensive consultation - existing literature - in-depth discussion	- reports submitted by institutes - Oasis database	- new/existing studies - meetings with managers - structured interviews	- questionnaire - bibliometrics - final/interim report
Evaluation Performance	Evaluator	- external > internal	- external > internal	- external > internal	- external < internal	- external < internal
	Method	- literature review - extensive interview	- public consultation (3 times) - literature review	- reports review - bibliometrics - site visit	- field survey - literature review - in-depth interview	- review of final report - questionnaire analysis - bibliometric analysis
	Criteria/ Indicator	- quantitative indicator • trained scientists etc. - qualitative indicator • new knowledge etc.	- appropriateness of organisation (stage 1) - effective./efficiency of management(stage 2)	- quantitative indicator • No. of articles etc. - qualitative criteria • level of KT etc.	- rationale & relevance - additionality - principal beneficiaries - appropriateness etc.	- quantitative indicator • No. of papers etc. - qualitative criteria • aim achievement etc.
Evaluation Utilisation	Report	- open to the public - posted on the website	- published - posted on the website	- not open to public - summary on website	- published on website - CD-ROM, hard copy	- not open to public - internal circulation
	Utilisation Type	- fully reflected in SR 2000 settlement - main inputs to S&I White Paper 2000	- inputs to SR 2002 - basic materials for developments for RCs	- budget allocation - development for institutes - inputs to BBSRC's 10-year road map & 5-year research plan	- mostly accepted - to be reflected in policies and strategies of civil space activity	- decision for further grant - identifying priority areas for future funding

# **CHAPTER 5. CASE STUDY OF THE KOREAN NATIONAL R&D EVALUATION SYSTEM**

## **5. 1. Introduction**

In recent times, the Korean government has made many efforts to enhance the national R&D evaluation systems, as a part of diverse policy means adopted for the improvement of national R&D activities. Although these efforts have initially resulted from the independent aims of the development of management systems in large-scale R&D programmes, they have also been affected by the introduction of legal provisions for the policy evaluation at the national level in 2001. As a result, it seems that Korea is one of the developing countries which have advanced R&D evaluation systems and have also conducted R&D evaluation actively and regularly (OECD, 2000 : 6).

In this chapter, using the same methodology and research structure for the UK case study, I shall conduct a case study of Korean R&D evaluation systems at the national level. This chapter comprises two main sections excluding the introduction : Section 5.2 relates to the current state of the R&D evaluation systems, Section 5.3 deals with the recent evaluation exercises in the S&T and R&D sectors. In more detail, Section 5.2 includes broad and, where available, specific information which describes both the past and current state of the Korean R&D evaluation systems. It also covers more detailed issues such as government S&T organisations, a brief history of the development of R&D evaluation, the current hierarchy of national R&D evaluation systems and major referential statutes for R&D evaluation, including laws, presidential decrees and ministerial ordinances. Section 5.3 includes five recent evaluation practices, one in relation to science and technology policy, two in relation to research institutions and one each in relation to R&D programmes and R&D projects.

For the work in this chapter, I have firstly reviewed the existing research articles and reports including evaluation reports, as well as many documents published or posted on the websites of the organisations concerned. I also conducted interviews with previous

and current staff members of the related government organisations concerned, in order to complement what was unclear or unconfirmed in the literature review. Interviews were mainly conducted by telephone and e-mail, depending on the specific interview situations.

## **5. 2. Current State of the R&D Evaluation Systems in Korea**

In this section, in order to provide a general understanding of science and technology activities in Korea as a whole, I shall first describe the government organisations in the S&T sector. Next, I shall briefly explain the history of the developments of policy evaluation and R&D evaluation in central government. Finally, an account of the hierarchy of the national R&D evaluation systems in Korea will be given, alongside the prescription of both some major referential statutes in relation to R&D evaluation systems and recent evaluation exercises.

### ***5.2.1. Government Organisations for S&T***

There are a variety of different ways to approach this subject. However, following the overall Korean governance system in the S&T and R&D sector, I intend to suggest a simplified typology of the government organisations, in accordance with their main roles and missions in the S&T policy process,<sup>1</sup> even though they are sometimes slightly overlap or are duplicated by other organisations.

#### ***5.2.1.1. Overview of the Governance System of S&T***

In Korea, since the late 1990s, there have been crucial changes in both the policy-making mechanism and the management system in the areas of S&T and R&D activities. Regarding the policy-making mechanism, it has been argued that the creation of the National Science and Technology Council (NSTC) and the designation of the Ministry of Science and Technology (MOST) as the secretariat of the NSTC in 1999 were turning points in the S&T sectors. The NSTC was created to improve the efficiency of

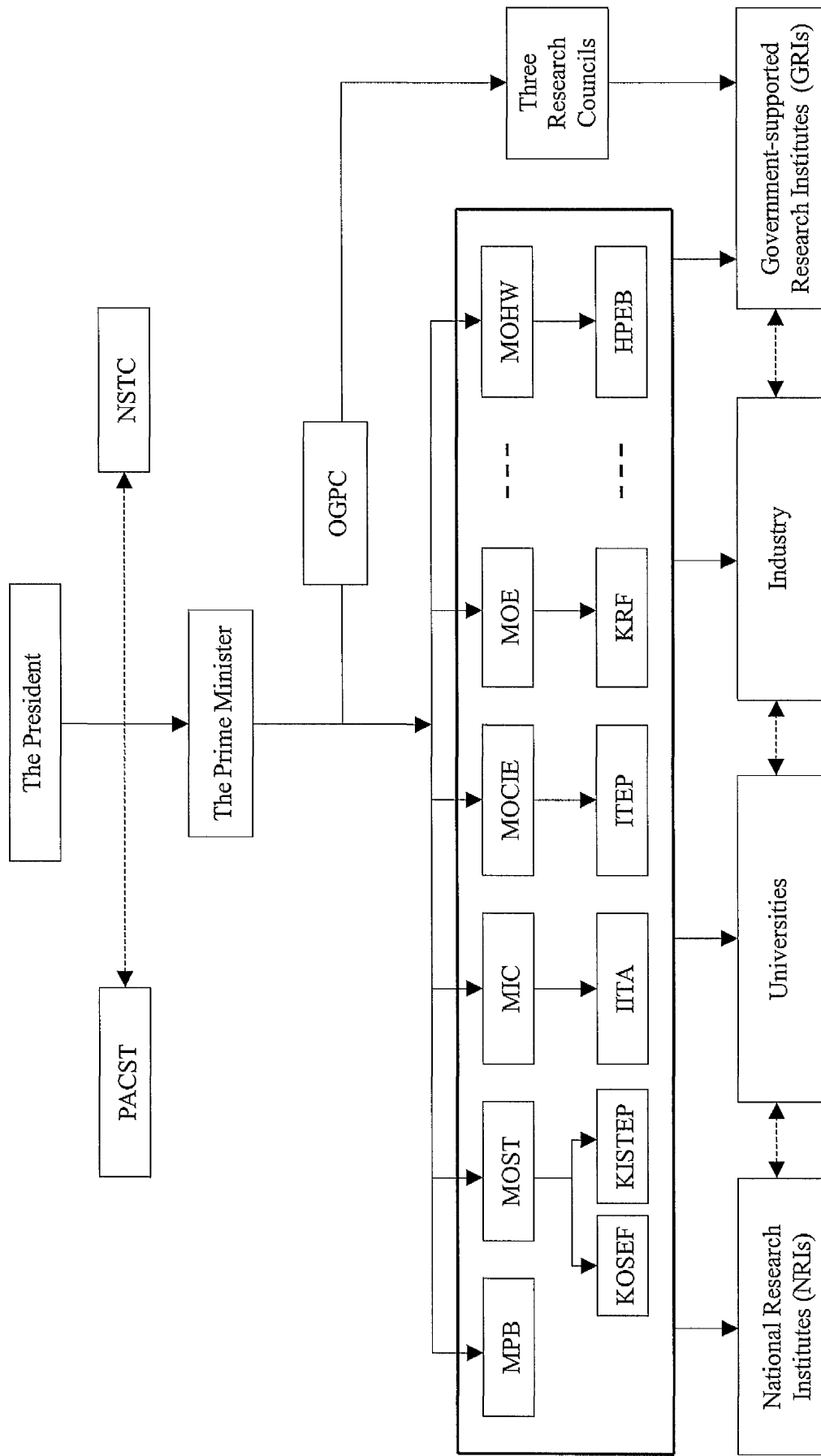
---

<sup>1</sup> With reference to the definition and coverage of the policy process, I have already given a brief explanation in Chapter 4 : The UK Case Study. For further information, see page 78.

government R&D activities through the inter-ministerial co-ordination of R&D policy and investment. As a result, the NSTC has the crucial mission of evaluating all national R&D programmes funded from government budgets as well as from public funds by the civil ministries. There has been tremendous criticism of the duplication and overlap that exist among public sector R&D programmes because of the lack of inter-ministerial co-ordination. For effective co-ordination, the NSTC is presided over by the President. In particular, in accordance with its mission, the MOST as the secretariat of the NSTC upgraded its seniority within the Cabinet, so that it can also function as a central co-ordinating body regarding S&T and R&D within the government (OECD, 2000 : 1).

Regarding the management system of R&D activities, reorganisation of the government-supported research institutes (GRIs) is considered one of the most important recent policy changes. This resulted from criticisms that the GRIs were not working in interests of the nation but at those of individual ministries to which they were answerable, and that the lack of inter-institutional mobility of R&D resources, for example, flows of human resource and information, and sharing of equipment, etc., among the GRIs caused the inefficient use of scarce resources. Therefore, most of the GRIs were regrouped into three research councils and put under the jurisdiction of the Office of the Prime Minister (OPM) in 1999. Three research councils are : (1) The Korea Research Council of Fundamental Science and Technology (KRCF), (2) The Korea Research Council of Public Science and Technology (KORP) and (3) The Korea Research Council for Industrial Science and Technology (KOCI). The chairperson of the research council reports directly to the Office of the Prime Minister, in particular, to the Office for Government Policy Co-ordination (OGPC) (OECD, 2000 : 2).

However, in this part, the emphasis will be placed not on the current policy changes but on providing an outline of the governance system of S&T, concentrating on the policy process and, in particular, policy formulation and policy evaluation. Therefore, my main purpose is to provide an account of the overall organisational structure of government S&T activity, particularly in terms of the responsibility for policy formulation, budget allocation and the performance of the R&D programmes. The principal Korean government components in the S&T sectors are shown in (Figure 5-1).



(Figure 5-1) Overview of the Korean System of Governance of S&T

In Korea, as in the UK case, S&T activities are categorised into two types. One type is basic and strategic S&T for future national competitiveness and the improvement of citizen's welfare as a whole, which is mainly funded by the MOST and the Ministry of Education and Human Resources Development (MOE). The other type is mission-oriented S&T for policy purposes which has been funded by civil ministries such as the Ministry of Commerce, Industry and Energy (MOCIE), the Ministry of Information and Communication (MIC), the Ministry of Health and Welfare (MOHW), the Ministry of Construction and Transportation (MOCT), the Ministry of Agriculture and Forestry (MAF) and so on. Of the ministries responsible for mission-oriented S&T activities, the MOCIE and the MIC are the most important supporters in terms of the amount of their investment in R&D programmes as well as of their past performance in S&T sectors. Consequently, R&D investment by the MOST, the MOE, the MOCIE and the MIC constitutes approximately 90 per cent of government civilian R&D expenditure (MOST, 2002a). However, there are slight differences between the policy process of the two types of S&T activities, which depend on the principal policy purposes emphasised in each circumstance.

Having given this background, I shall now give a brief description of government organisations and their roles in the area of S&T in Korea. Government organisations involved in S&T can be broadly grouped, according to their major roles in the S&T policy process, into those whose purpose is policy formulation, those who deal with budget allocation, and major R&D performers. These organisations are summarised in <Table 5-1>.

As already indicated in Chapter 1, regarding the Korean case study, the research scope of my thesis is focused on the evaluation systems and the current exercises which the NSTC, the Office of the Prime Minister/the Office for Government Policy Co-ordination, the MOST, the research councils (in particular, the KOCI) and the "R&D management institutes" (in particular, the KISTEP) have been operating, so my exploration of this subject will be mainly devoted to organisations related to the area of basic and strategic S&T activities.

**<Table 5-1> Typology of Government Organisations for S&T in Korea**

Policy Process S&T Type	Policy Formulation		Budget Allocation	Major R&D Performers
	Central Government Level	Ministry Level		
<b>Basic &amp; Strategic S&amp;T</b>	- NSTC/MOST - OPM/OGPC - MPB	- MOST - MOE	- MPB - MOST/KISTEP	- NRIs - GRIs - Universities
<b>Mission-Oriented S&amp;T</b>	- Research Councils	- MPB - MOCIE, MIC, MOHW, MOCT, MAF, etc.		- GRIs - GIIs - Universities

\* This is schematic for analytical purposes. In reality, the distinctions between basic and mission-oriented S&T are blurred.

### **5.2.1.2. Organisations for Policy Formulation**

The organisations involved in the formulation of policy for S&T are again divided into two types. While one includes “central government organisations”, such as the NSTC/the MOST, the OPM/the OGPC, the Ministry of Planning and Budget (MPB) and the research councils, the other includes “ministry organisations”, such as the MOST, the MOE, the MOCIE, the MIC and so on. The central government organisations set up a general and basic framework for S&T policy through various mechanisms, including budget allocation, the enactment of statutes and the issues of guidelines that affect both basic and strategic S&T activities and mission-oriented S&T activities. However, in principle, they do not prescribe in detail how it is formulated and implemented. The execution of detailed policy processes, in particular, at the stage of the policy implementation, are almost handed over to the individual ministries directly concerned. The ministry organisations have a mission to formulate more detailed and specific S&T policy for their policy purposes. These activities normally cover the articulation of their needs for S&T activities, the generation and management of R&D programmes and projects, and the promotion of the relationship with customers in their domain in line with the guidelines or advice of the central government organisations.

#### **(1) Central Government Organisations**

Regarding this subject, in principle, it is considered that the NSTC has the highest-level responsibility for S&T activities in Korea. However, the MOST has the actual power to



handle S&T affairs at the national level, through the mission of the secretariat of the NSTC, because the NSTC is a non-standing committee. In addition, the OPM/the OGPC, the MPB and the research councils act as central government organisations in some areas of S&T-related policy issues.

**National Science and Technology Council / Ministry of Science and Technology** : As already noted, in Korea, there has been tremendous criticism of the duplication and overlap that exist among the S&T activities in the public sector because of the lack of inter-ministerial co-ordination. So, in order to review and co-ordinate national S&T policies and R&D programmes and to set the priority for the allocation of R&D budgets at the national level, the Korean government established the NSTC, whose mission was expanded in January 1999 (OECD, 2000 : 1 ; MOST, 2002b).

As of January 2002, the NSTC was chaired by the President and comprises 15 ministers of S&T-related ministries and nine representatives from the S&T community including industry, academia and the research institutes. Of these members, the Minister of the MOST acts as the executive secretary of the NSTC and the MOST currently serves as its secretariat. The committee has three meetings a year, which are normally held in April, July and December. Meanwhile, under the NSTC, there are three sub-committees and four expert committees who deal with some specific S&T-related subjects. The three sub-committees are the Operating Committee of the NSTC, the Committee of the Bio-technology and Bio-industry and the Committee of the Regional S&T Promotion, and the four expert committees include the Expert Committee for S&T policy, the Expert Committee for R&D, the Expert Committee for Aerospace Development and the Expert Committee for Nano-technology (NSTC, 2002a ; MOST, 2002b).

As the secretariat of the NSTC, the MOST currently plays the role of a central agency overseeing national S&T policy. It has administered S&T affairs at the national level and co-ordinated national R&D programmes since it was established in 1967. The main functions of the MOST include : (1) formulating policies for S&T development and to perform technology forecasting, (2) developing core technology, future-oriented technology and large-scale technology, (3) pursuing technological self-reliance for the safe use of nuclear technology, (4) supporting basic and applied research conducted by the GRIs, universities and private research institutes, (5) formulating policies for R&D

investment, human resources development, S&T information and international S&T co-operation and (6) promoting public awareness of S&T. Meanwhile, recently, the Korean government, recognising that S&T is essential for Korea's transition towards being an advanced country, has upgraded the position of the MOST in the Cabinet. As a result, in practice, the MOST is fully responsible for co-ordinating S&T policies among the ministries concerned and for overseeing their compliance with related policies (MOST, 2002b).

**Office of the Prime Minister / Office for Government Policy Coordination** : Under the Office of the Prime Minister (OPM), there are 11 administrative offices which deal with certain policy missions. Of these, to assist the Prime Minister, the Office for Government Policy Coordination (OGPC) is in charge of supervision, co-ordination, policy analysis and evaluation, and the deregulation of the administration of each central government organisation. Then, there are also five Assistant Ministers for “the Public Policy Co-ordination”, “the Economic Policy Co-ordination”, “the Social and Cultural Policy Co-ordination”, “the Policy and Evaluation” and “the Regulatory Reform” in the OGPC. Of these, “the Assistant Minister for the Policy Analysis and Evaluation” is fully engaged in the policy evaluation of both the central and local government performance, through acting as the secretariat of the “Policy Evaluation Committee” which is answerable to the Prime Minister, whereas “the Assistant Minister of the Regulatory Reform” currently supervises and supports three S&T-related research councils and their 19 GRIs (OPM, 2002a).

Consequently, the OGPC is fully accountable for the evaluation of policy, including S&T and R&D policy, at the national level. It also shares some responsibilities for R&D matters with other central government organisations such as the MOST and the MPB, in particular, through the supervision of the research councils and their affiliated GRIs.

**Ministry of Planning and Budget** : The MPB, placed in the Cabinet under the Prime Minister, is the main agency in the Korean government responsible for national budget-related policy, budget formulation and execution management, the planning and management of public funds, and the reform of the government and public sector, including government funded enterprises and other government supported institutions (MPB, 2002a).

In relation to S&T affairs, the MPB exercises substantial influence over the ministries concerned through budgetary control. At the start of the fiscal year, the MPB allocates an S&T budget to ministries on a lump sum or a project basis. Each ministry then decides the detailed distribution of the budgetary package on the basis of its programmes. Since ministries compete for funds, the budgetary control function of the MPB requires both co-ordination and authority (OECD, 1996 : 33). As a result, in general, the MPB carries out the crucial function of guiding the principal direction and structure of the national S&T activities.

**The Research Councils** : As already noted, in Korea, there are three S&T-related research councils under the Office of the Prime Minister. These generally supervise and guide the R&D activities of the GRIs, who are the main performers in the basic and strategic S&T sector.<sup>2</sup> In particular, research councils are in charge of the structural reinforcement of the autonomy and responsibility of the GRIs and of eliminating the interference of other ministries and/or offices at the national level. Meanwhile, the Korean research councils are not government ministries or agencies but semi-government organisations,<sup>3</sup> although they actually function as central government organisations.

In general, the missions of the research councils are to support, foster and systematically manage the GRIs under their controls, and thus undertake to assist and promote scientific and technical research and knowledge-based industries in various sectors of national importance. To carry out these missions, each individual research council undertakes the following detailed functions : (1) to provide research planning in the area of its responsible S&T and to co-ordinate future strategies for the GRIs under its supervision, (2) to review and improve functions of its GRIs and to ensure their specialisation in unique, pioneering science and technology, (3) to establish an innovative, knowledge-based evaluation system for efficient R&D activities and autonomous research management and (4) to ensure sustainable funding resources for carrying out focused collaborative research among the GRIs (KRCF, 2002 ; KORP, 2002 ; KOCI,

---

<sup>2</sup> In addition, there are two research councils that supervise GRIs in the areas of social sciences and humanities. These are the Korea Council of Economic and Social Research Institutes (KCESRI) and the Korea Council of Humanities and Social Research Institutes (KCHRIS), under which there are currently 14 GRIs and nine GRIs, respectively.

<sup>3</sup> The Korean research councils were established as one type of the government-supported institutions in 1999, based on the Act on Establishment, Operation and Promotion of the Government-supported Research Institutes (GRIs). Therefore, in principle, they have more autonomy from related ministerial controls on

2002a).

Therefore, in fact, the three research councils act as one of the central government organisations in the areas of basic and strategic S&T activities, in collaboration with other government organisations such as the MOST, the OGPC and MPB, and sometimes in line with their guidance and advice.

## **(2) Ministry Organisations**

At the ministry level, the MOST and the MOE are mainly involved in the basic and strategic R&D activities which aim to upgrade general S&T capabilities at a national level. Thus, the MOST normally has a double position in the policy formulation systems for S&T and R&D activities, in that, while it acts as the secretariat of the NSTC at the national level, it also conducts large-scale national R&D programmes for its own organisational missions. As the former has already been explained, I shall focus here on the latter.

**Ministry of Science and Technology** : Initially, the role of the MOST was to develop a strategy for the rapid development of the Korean science and technology. To carry out its mission, it has planned and conducted various types of national R&D programmes since 1982. Consequently, it acts as a main policy formulator in the basic and strategic S&T sector, because the national R&D programmes normally contain various research communities such as the GRIs, universities and private companies.

The national R&D programme of the MOST was expanded and diversified during the 1980s and 1990s, in response to the change in both the national needs for S&T activities and in the policy circumstances of the public sector. Recently, the MOST's national R&D programme has placed greater emphasis on emerging technologies, such as biotechnology, IT technology, nanotechnology and aeronautics, as well as on conventional industrial technologies, such as textiles and shipbuilding. The national R&D programme is divided into more detailed sub R&D programmes, according to the specific research objectives. These include (1) the 21st Century Frontier R&D Programmes, (2) the Highly Advanced National Project (HAN Project), (3) the Creative Research Initiatives (CRI), (4) the National Research Laboratory (NRL) Programme, (5)

---

issues relating to day-to-day activities but remain accountable to the Minister of the OGPC.

the Biotechnology Development Programme and (6) the Space and Aeronautics Programme. In 2001, the MOST invested a total of US\$ 400 million in these R&D programmes and projects (MOST, 2002c).

Meanwhile, at the stage of policy implementation, detailed policy means, including the allocation of the research grants and the evaluation of the R&D projects, have been handed over to the Korea Institute of Science and Technology Evaluation and Planning (KISTEP) and the Korea Science and Engineering Foundation (KOSEF), its affiliated institutions, which are responsible for national R&D management and for the support of basic science, respectively.

**Ministry of Education and Human Resources Development (MOE)** : The MOE has jurisdiction over affairs relating to higher education in the university, except for the Korea Advanced Institute of Science and Technology (KAIST) and the Kwang-ju Institute of Science and Technology (K-JIST), which are research focused higher education institutes founded by the MOST. The MOE also has responsibility for Korea's entire education system, including primary and secondary schools, adult education and university arts, humanities and social science programmes, in addition to university physical science and engineering. Traditionally, the MOE's attention has primarily focused on primary and high school education, and on teaching rather than on research in the universities, whereas the MOST is mainly concerned with the promotion of scientific achievement and excellence in universities (OECD, 1996 : 162).

However, in order to improve research quality in the universities, the MOE has continued to implement a series of policies. Of these, it seems that the "Brain Korea 21" (BK 21) programme in 1999 has been considered as a turning point in the research system in the universities. This programme has aimed at fostering world class scholars in research to found the creative and advanced knowledge base necessary for the 21st century. The major purposes of the BK 21 programme are : (1) to foster world-class research universities which will serve as an infrastructure to produce ideas and technology that are creative and original, (2) to strengthen the competitiveness of local universities, (3) to introduce professional graduate schools to train professionals in their research field and (4) to create an environment where universities compete with each other not based on their prestige, but based on the quality of research outcomes and student performance. For

these purposes, the government will provide intensive support for professors and graduate students so that they can concentrate on research activities. It will also invest a total of US\$ 1.2 billion over the seven years from 1999 (MOE, 2002a).

### ***5.2.1.3. Organisations for Budget Allocation***

In Korea, as already mentioned, the MPB firstly allocates an S&T budget to the ministries, which carry out the basic and strategic R&D programmes as well as the mission-oriented R&D programmes. Then, related ministries use their judgement to re-allocate them to the GRIs, universities and industries, via their “R&D management institutes” which deal with R&D programmes conducted for their policy purposes. At present, there are nine R&D management institutes in Korea. They include the KISTEP of the MOST for basic and strategic S&T, the Korea Institute of Industrial Technology Evaluation and Planning (ITEP) of the MOCIE for industrial technology, the Institute of Information Technology Assessment (IITA) of the MIC for IT technology and the Health Technology Planning and Evaluation Board (HPEB) of the MOHW for health technology (MOST, 2002h). Regarding the budget allocation in the S&T sector, the MPB plays a major role at the national level, whereas the R&D management institutes such as the KISTEP, the ITEP, the IITA and the HPEB are main actors at the level of each ministry concerned. In contrast to the UK case, the Korean research councils have no power to allocate research funds to the GRIs, universities or industries.

As a result, in Korea, both the MPB and each R&D management institute under individual ministry have the principal responsibility for the allocation of research funds, even though the level of the policies and programmes, which are emphasised by each budget organisation, is greatly different. Regarding the R&D management institutes, I shall here deal only with the KISTEP, in accordance with the research scope and level described in Chapter 1. Also, the MPB will be excluded, in order to avoid explaining it again, because it has already been mentioned above.

#### **(1) The Korea Institute of S&T Evaluation and Planning (KISTEP)**

Since its foundation in February 1999, the KISTEP has carried out three major functions :

(1) to assist the NSTC in conducting surveys, analyses and evaluations of the national R&D programmes, which are implemented by several ministries, including the MOST,

(2) to work for the MOST in planning, managing and evaluating the national R&D programmes sponsored by that ministry and (3) to implement international S&T co-operation programmes to promote the globalisation of the nation's science and technology. In recent, the KISTEP has also made a new start through expanded missions such as providing expertise and administrative assistance to the secretariat of the NSTC, namely the MOST, in conducting pre co-ordination of the national R&D programmes as well as undertaking technology assessments and technology foresight (KISTEP, 2002a).

Of three missions of the KISTEP, the second role is directly related to the budget allocation to the research performers such as the GRIs, universities and industries, in particular, in the area of basic and strategic R&D programmes. That is, the "Division of National R&D Planning and Management", one of the four divisions and one centre in the KISTEP, manages the national R&D programmes and projects of the MOST, in order to enhance national competitiveness and the quality of life in the mid and long-term perspectives. More specifically, the division selects and contracts R&D projects in line with government S&T policies, particularly those of the MOST (KISTEP, 2002a). Through this work, the KISTEP acts as a main budget distributor to the research communities.

#### ***5.2.1.4. Organisations for R&D Performance***

In Korea, three types of public research institutes have conducted related programmes in the area of basic and strategic research : the national/public research institutes (NRIs), the government-invested institutions (GIIs) and the government-supported research institutes (GRIs). Among these, the GRIs have captured the lion's share in term of investment, manpower and facilities (OECD, 1996 : 63 ; OECD, 2000 : 4). Recently, the Korean government has endeavoured to promote basic science and research in the universities that, despite an abundant brain pool in Korea, have shown relatively poor performance because of insufficient research funds and facilities.

As a result, in Korea, the GRIs and universities have played a major role in carrying out the national S&T and R&D programmes in the area of basic and strategic research. They receive most research funds for R&D from the government budget and public funds (e.g.

the Informatisation Promotion Fund of the MIC) on a competitive basis. In addition, they also receive research funds from the related government ministries (e.g. the MOCIE, the MOHW, the MOCT and the MAF etc.) and industries.

### **(1) Government-supported Research Institutes (GRIs)**

The Korean GRIs have a very special position in the national innovation system in that they have carried out long-term and strategic R&D programmes and made a great contribution to economic growth and societal development. That is, they have assumed the leading role in the S&T and R&D activities in the public sector. They are semi-government organisations. Even though they are paid by the government, the researchers of the GRIs are not civil servants (OECD, 1996 : 63).

In the late 1960s, Korea lacked the technological capability necessary for industrialisation. So, the Korea Institute of Science and Technology (KIST) was established in 1966 as an integrated technical centre to meet industrial needs. In the 1970s, the expansion of the Korean economy and the great sophistication and diversity of its industrial structure required that both the volume and scope of public R&D keep pace with changes. A number of specialised research institutes were established as spin-offs from the KIST. Each was designed to develop in-depth capabilities in strategic areas : namely, shipbuilding, marine resources, electronics, telecommunications, energy, machinery and chemicals. In the 1980s, the government consolidated the 15 GRIs that had been responsible for various ministries into nine large GRIs under the MOST (OECD, 1996 : 64-67 ; Yi and Kang, 2000b : 160-161).

During the late 1990s, for the reasons already noted, most GRIs had to redefine their roles in the national innovation system. In order to sort this matter out, the Korean government enacted a new law in January 1999, known as "The Act on the Establishment, Operation and Promotion of the Government-supported Research Institutes (GRIs)". According to this law, a new system for the management and promotion of the GRIs has been created since 1999. Under the new system, which took the UK and the German experiences as a benchmark in public R&D activities, there are three S&T related research councils, the KORP, the KRCF and the KOCI, under the Office of the PM, which supervise their affiliated GRIs. This new system endowed individual GRIs with more autonomy in operation, management, decision-making and organisational change. At the same time,



the GRIs have to be more accountable for the performance of the R&D projects and of their management system. Consequently, the new system is expected to improve research productivity, to strengthen the linkage between research institutes and to increase knowledge transfer and the commercialisation of research outcomes. Despite this system, eight GRIs still remain under the MOST umbrella (Yi and Kang, 2000b : 161 ; MOST, 2002e).

As of December 2001, there were 27 GRIs under the three research councils and the MOST. They employed about 10,000 research staff, of whom about 7,000 belonged to the GRIs under the research councils and approximately 3,000 to the MOST. Also, the 27 GRIs spent R&D funds of KRW (Korean Won) 2,096.5 billion, of which 29 % came from government budget and 71 % from other research funds such as public funds and industries. The current state of the Korean GRIs, including research staff and R&D expenditures, is shown in <Table 5-2> (MOST, 2002a : 265-267 ).

**<Table 5-2> Current State of the GRIs in Korea (2001)**

<b>Name of GRI</b>	<b>Research Field</b>	<b>Staff</b>	<b>Budget (KRW million)</b>
<b>The Korea Research Council of Fundamental Science and Technology (KRCF)</b>			
Korea Basic Science Institute (KBSI)	R&D Facility Support	171	60,563
Korea Astronomy Observatory (KAO)	Astronomy	79	10,225
Korea Institute of Science and Technology (KIST)	National Projects	652	114,596
Korea Research Institute of Bioscience & Biothechnology (KRIBB)	Bio-science	242	53,244
<i>Sub Total</i>	<i>4 GRIs</i>	<i>1,144</i>	<i>238,628</i>
<b>The Korea Research Council of Public Science and Technology (KORP)</b>			
Korea Institute of Science and Technology Information (KISTI)	S&T Information Service	276	61,090
Korea Institute of Construction Technology (KICT)	Construction	231	35,166
Korea Railroad Research Institute (KRRI)	Railroad	179	31,409
Korea Ocean Research & Development Institute (KORDI)	Ocean	332	68,908
Korea Research Institute of Standards & Science (KRISS)	National Standards	330	63,779
Korea Institute of Energy Research (KIER)	Energy	277	53,101
Korea Institute of Geoscience and Mineral Resources (KIGMM)	Geoscience	339	51,189
Korea Aerospace Research Institute (KARI)	Aerospace	325	89,530
<i>Sub Total</i>	<i>8 GRIs</i>	<i>2,289</i>	<i>454,172</i>

<b>The Korea Research Council for Industrial Science and Technology (KOCT)</b>			
Korea Institute of Oriental Medicine(KIOM)	Oriental Medicine	30	3,594
Korea Institute of Industrial Technology (KIT)	Industrial Technology	224	59,882
Electronics and Telecommunications Research Institute (ETRI)	IT Technology	1,976	349,691
Korea Food Research Institute (KFRI)	Food	141	16,749
Korea Research Institute of Machinery & Materials (KIMM)	Machinery and Materials	417	67,869
Korea Institute of Chemical Technology (KRIT)	Chemistry	359	70,355
Korea Electrotechnology Research Institute (KERI)	Electrotechnology	279	61,982
<i>Sub Total</i>	<i>7 GRIs</i>	<i>3,426</i>	<i>630,032</i>
<b>The Ministry of Science and Technology (MOST)</b>			
Korea Advanced Institute of Science and Technology (KAIST)	S&T Education	681	173,126
Kwang-ju Institute of Science and Technology (K-JIST)	S&T Education	130	39,513
Korea Institute of Advanced Study (KIAS)	Basic Science	72	5,954
Korea Science and Engineering Foundation (KOSF)	Support to Basic Science	98	230,688
Korea Atomic Energy Research Institute (KAERI)	Atomic Energy	1,028	180,716
Korea Institute of Nuclear Safety (KINS)	Nuclear Safety	300	38,461
Korea Cancer Centre Hospital (KCCH)	Cancer & Radiation Medicine	738	86,910
Korea Institute of S&T Evaluation and Planning (KISTEP)	S&T Evaluation and Planning	100	18,523
<i>Sub Total</i>	<i>8 GRIs</i>	<i>3,147</i>	<i>773,891</i>
<b>Total</b>	<b>27 GRIs</b>	<b>10,006</b>	<b>2,096,543</b>

Source : MOST (2002a)

## (2) Universities

As in other countries, in Korea, universities are also main performers in the area of basic research. They represent the biggest brain pools consisting of 78 % of all current Ph.D. researchers. However, the research conditions in universities are not optimal due to the lack of funds, research equipment and other facilities. Faculties also suffer from teaching overloads, which have been considered as another reason for the low R&D performance. In 2000, the total amount of national R&D expenditures in Korea was KRW 13,848.5 billion (about US\$ 12,249 million). Only 11.3%, namely KRW 1,561.9 billion, of this was allocated to universities, whereas the figure for public research institutes including the GRIs was 14.76% and that for industry was 74.0%. While 60.4% (KRW 943.1 billion)

of total R&D expenditures by the universities came from the government and the public sector, the figure from the private sector was 39.4% (KRW 615.2 billion) and that from overseas was 0.2% (KRW 3.5 billion) (MOST, 2002a : 402 ; 430). As of December 2001, in Korea, there were 162 universities with postgraduate programmes, consisting of 24 national universities, two public universities and 136 private universities (KEDI, 2002).

As already noted, the Korean government's support for basic research is funded mainly by the MOST and the MOE via their respective agencies, the Korea Science and Engineering Foundation (KOSEF) and the Korea Research Foundation (KRF). In particular, the MOST has supported R&D activities in universities through the Centres of Excellence (COE) programme via the KOSEF. There are three types of COE in Korea : the Science Research Centres (SRCs), the Engineering Research Centres (ERCs) and the Regional Research Centres (RRCs). These COEs were set up to implement programmes that encourage basic research in major universities. While the SRCs and ERCs, founded in May 1989, focus on creative and innovative research in basic science and new technologies, the RRCs, started in 1995, emphasise co-operative research between regional universities and industries. The SRCs and the ERCs were selected on the basis of creativity and research capability. In the selection of RRCs, both research capability and contribution to the regional economy and community are important factors. Once the centres are selected, they receive government funding for nine years providing that the interim evaluation conducted every three years shows good progress. Up till December 2001, a total of 135 COEs, comprising 39 SRCs, 51 ERCs and 45 RRCs, have been selected and funded. The number of these COEs is expected to increase to 150 by 2002 (MOST, 2002a : 84-87 ; MOST, 2002b).

Meanwhile, the MOE initiated the BK-21 programme in 1999 via the KRF, in order to upgrade the quality of university research and postgraduate education. The focus of this programme is on advanced S&T areas such as IT, biotechnology and nanotechnology. The BK-21 programme provides selected universities with a fund of KRW 200 billion every year for seven years, which has to be spent supporting postgraduate programmes with excellence in advanced research and education. Therefore, the beneficiaries of this programme have to reduce their undergraduate enrolments and concentrate their resources more on postgraduate programmes and research activities. (OECD, 2000 : 2-4).

### ***5.2.2. Brief History of the Development of R&D Evaluation***

In Korea, in contrast to the UK, where the policy evaluation and the R&D evaluation systems have been developed with a close linkage between them, two evaluation systems have been evolved separately to meet their own requirements. It is only very recently that some relationships have been developed between them, as a part of efforts by the Korean government to connect and, where appropriate, integrate all the existing evaluation systems, including R&D evaluation system.

Hence, in this part, I shall firstly provide a brief history of the two evaluation systems in Korea. Then, a brief explanation will be given of the relationships between policy evaluation and R&D evaluation.

#### ***5.2.2.1. Development of Policy Evaluation***

In Korea, the evaluation system for the government performance, including the outputs of the main policies, was introduced for the first time in 1961. Although it has a relatively long history, it appears that, until the late 1990s, its importance and roles had not been properly recognised among the high-level policy makers so that the institutional development necessary for policy evaluation had not progressed appropriately in response to the changes in the political and administrative environment (Cha, 1998 : 38). In what follows, the process of development of policy evaluation system will be described in brief, on the basis of information acquired from the existing literature (Kim, 1993 ; Lee and Suh, 1996 ; Cha, 1998 ; PEC, 2002a). This description indicates how the development stage in the Korean policy evaluation system is based upon the extent to which the evaluation system has been institutionalised and to which the evaluation practice has been activated.

##### **(1) Introduction Stage (September 1961 ~ October 1981)**

The first time that the Korean government conducted a policy evaluation was in 1961, as a part of the evaluation of government performance, which was considered as a more or less comprehensive evaluation type. During this period, the Planning and Co-ordination Bureau under the Office of the Prime Minister (OPM) evaluated government policies.

Every year, a basic management plan for overall government duties was drawn up and followed by an evaluation and analysis of the results of implementation.

## **(2) Transitional Stage (November 1981 ~ December 1994)**

At this period, it is argued that evaluation functions were weakened, compared with those of the previous period. At this time, the responsibilities for policy evaluation were assigned to the two different central offices : the Economic Planning Board (EPB) and the Administrative Co-ordination Office under the Office of the Prime Minister. While the Economic Planning Board carried out the *ex-post* policy evaluation as well as the *ex-ante* policy analysis of government policies, the Administrative Co-ordination Office focused more on policy evaluation activities.

In 1981, the responsibility for the evaluation and analysis of the OPM were transferred to the Policy Evaluation Bureau of the EPB. Individual ministries and agencies were asked to analyse their own duties, while the Economic Planning Board concentrated on collecting and co-ordinating analysis results from each central administrative office. Major government policies and programmes were evaluated, but the focus was placed on the analysis of the progress of programmes and on the evaluation of management systems in government-invested institutions.

Meanwhile, in 1990, with the assignment of the function of policy evaluation to the Administrative Co-ordination Office (ACO) of the OPM, it carried out a complementary role in the analysis system of the EPB, which was limited to diagnosing overall government activities. There were differences between the policy evaluation of the ACO and the policy analysis of the EPB. In the ACO's policy evaluations, the progress of major policies in overall government was checked and problems pinpointed. By reflecting these results in each central administrative office's plans, the effectiveness of government policies could be enhanced. In the EPB's policy analysis, the progress and the causes of problems in individual and budgeted programmes were analysed in order to push for the reasonable implementation of the programmes in question.

## **(3) Readjustment Stage (December 1994 ~ February 1998)**

Following the reorganisation of the government structure in December 1994, the government performance evaluation system was revamped by assigning all policy

analysis duties, which had been under the EPB, to the Administrative Co-ordination Office under the Prime Minister. Alongside the integration of policy evaluation with policy analysis, a more comprehensive and systematic review of the evaluation system was carried out.

The purpose of the transfer of the responsibilities for policy evaluation was to enhance the efficiency of government policies and programmes and to secure accountability for government policies. Evaluation activities were divided into (1) regular evaluations for major policies selected for annual evaluation and (2) specific evaluations when deemed necessary in connection with the management of government affairs

#### **(4) Development Stage (March 1998 ~ Present)**

With the establishment of the Policy Evaluation Committee (PEC) in April 1998 and the enactment of the “Framework Act on Government Performance Evaluation” in January 2001,<sup>4</sup> there was systematic development of the policy evaluation system in Korea. With the change of the President in 1998, the new government started to revamp and reinforce a policy evaluation system. As a result, the current policy evaluation system in Korea has three characteristics such as : (1) the introduction of the “institution evaluation system”, (2) the establishment and operation of the Policy Evaluation Committee and (3) the enactment of the Framework Act on Government Performance Evaluation.

Firstly, the existing policy evaluation system was limited in its ability to conduct a evaluation of the comprehensive potential of each central administrative office because only a portion of the functions could be evaluated. It has also been pointed out that, because civil servants initiated the evaluations, objectivity and professionalism in evaluation could not be fully achieved. In order to make evaluations more effective and to ensure accountability, the Korean government introduced institution evaluation systems whose aim was to comprehensively evaluate the entire capabilities of a certain government ministry or agency, including main policy outputs and its willingness to carry these out.

As well as conducting evaluations of government performance as a whole, in order to

---

<sup>4</sup> With reference to three evaluation practices in 1998, 1999 and 2000, the government performance evaluation was conducted on the basis of not a law but a presidential decree.

enhance the efficiency and accountability of administration, the government introduced an incentive system in the public sector based on the evaluation findings. Also, evaluation standards and indicators to represent the specialities of each area of evaluation were established in order to make the evaluation credible. In addition, the feedback functions and evaluation utilisation in the evaluation process have been greatly strengthened.

Secondly, in order to secure the professionalism and objectivity of evaluations, under the Office of the Prime Minister, the Policy Evaluation Committee, which comprises around 30 civil experts from academia, the research institutes and industry, was established to participate actively in the evaluation process.

Thirdly, to establish an efficient evaluation system for government functions and to promote the efficiency and integrity of local governments by establishing evaluation standards, in November 1999, the Korean government accepted the Policy Evaluation Committee's proposal to legislate the Framework Act on Government Performance Evaluation. After public hearings, this act was passed in January 2001 and put into operation in May of the same year.

Consequently, through development process of the policy evaluation system, the Korean government is now conducting evaluation of the main policies of all central ministries and agencies at the national level, including the MOST, as a part of its government performance evaluation.

#### ***5.2.2.2. Development of R&D Evaluation***

Even though national R&D activities have had a key role in economic growth and societal development in Korea since 1960s, the evaluation of R&D programmes/projects and management system of research institutes was not activated until recently (Yi and Kang, 2000b : 161). On the other hand, in the R&D evaluation system at the national level, three types of R&D evaluation activity, namely the evaluation of R&D programme, the evaluation of R&D project and the evaluation of research institution have evolved separately in response to the changes in their situation. Therefore, there has been criticism of overlapping and duplication among R&D evaluation activities, and this may be considered as one cause for a lot of administrative burdens on researchers and research

institutes and, in particular, on the GRIs and their research staff.

In Korea, the first R&D evaluation activity that the MOST undertook was the evaluation of its supporting national R&D programme (called the Special R&D Programme) in the area of strategic technology, which was the first national R&D programme initiated by the government in 1982. At that time, the R&D programme was evaluated not at the level of the programme itself, but at that of individual projects composing it. During the period between 1982 and 1991, the MOST generally conducted evaluation activities in the area of all R&D projects, even though some evaluation tasks were delegated to the affiliated institutes concerned. This structure of the MOST's evaluations resulted in criticism of a lack of professionalism and objectivity in the evaluation practices (Kim, 1999 : 17).

Throughout the late 1980s and the early 1990s, in addition to the MOST's national R&D programme, diverse national R&D programmes were launched by other civil ministries in Korea. These programmes included the MOCIE's "Industrial Generic Technology Development Programme" in 1987, the Ministry of Energy and Resource (MOER)'s <sup>5</sup> the "Alternative Energy Development Programme" in 1988, the MIC's "Information and Communication Technology Development Programme" in 1992 and the Highly Advanced National Project (HAN Project) launched in 1992 by many ministries concerned, including the MOST, the MOCIE, the MIC and the MOHW. Alongside the establishment of lots of national R&D programmes in different areas of S&T, the related ministries set up so called the "R&D management institutes" under their control in order to enhance the effectiveness of the R&D management process. They also entrusted the mission of R&D management, including *ex-post* evaluation, to their affiliated institutes. These institutes included the Science and Technology Policy Institute (STEPI) <sup>6</sup> under the MOST, the ITEP under the MOCIE and the IITA under the MIC. As a result, individual institutes under the control of related ministries were in full charge of R&D project evaluation in each area of S&T. Although there was a great deal of development in the evaluation system in this period, evaluation practices were similar to those of the previous period. In other words, the R&D management institutes have evaluated their national

---

<sup>5</sup> This ministry was later merged into the MOCIE.

<sup>6</sup> The STEPI then comprised two main divisions : the S&T Policy Research Division and the R&D Management Division. In February 1999, the latter division was separated from the STEPI and expanded into a new institute known as the KISTEP, which is more independent and professional, on the basis of the "Special Act for Scientific and Technological Innovation".



R&D programmes not from the perspective of the nationwide S&T and R&D policies, but from each ministry's point of view. Furthermore, they have conducted their evaluations not at the level of the programmes themselves but at that of the projects (Yi and Kang, 2000b : 161).

For the above-mentioned reasons, R&D evaluation activities and, in particular R&D programme/project evaluation, in Korea have not achieved their purposes, which include the optimal allocation of R&D resources and the production of useful policy information for the improvement of R&D activity. In order to deal with this deficiency in R&D evaluation, the Korean government revised the "Special Act for Scientific and Technological Innovation"<sup>7</sup> in February 1998. On the basis of this law, in principle, the NSTC and, in practice, the MOST, as the secretariat of the NSTC, were authorised to evaluate all national R&D programmes from a nationwide perspective. In actual fact, the evaluation works of both the national R&D programmes and the national R&D projects of the MOST are delegated to the KISTEP, in accordance with the Framework Act on Science and Technology in 2001(MOST, 2001a).

Meanwhile, regarding the evaluation of research institutions, the MOST had conducted evaluations of only its affiliated GRIs between 1991 and 1997. However, as these were not performed due to any legal requirements, but due to the MOST's internal necessities, institution evaluation, for example, the evaluation of the GRIs and the NRIs, was not the general case in Korea (Lee and Suh, 1996 : 245). As already noted, in January 1999, in order to lay down the responsibility for the evaluation of research institutes in the R&D sector, the Korean government enacted the "The Act on the Establishment, Operation and Promotion of the Government-supported Research Institutes (GRIs)". It also gave three research councils the missions of evaluating the performance of R&D projects and the management systems of their affiliated GRIs. In addition, this law endowed the OGPC under the OPM and the NSTC with the responsibility for the evaluation of the research councils. At early stage, both the OGPC and the NSTC had each mission for the evaluation of five research councils. The former was responsible for the evaluation of two research councils in the area of social sciences and humanities, whereas the latter was charged with that of three ones in the area of science and technology, respectively.

---

<sup>7</sup> This law was replaced by a new law known as the "Framework Act on Science and Technology" in January 2001

However, in 2001, the evaluation work of the NSTC was amalgamated into the OGPC (OPM, 2001c). As a result, at the moment, the OGPC evaluates all five research councils collectively.

Therefore, having started to evaluate all national R&D programmes as well as the GRIs and the research councils mainly responsible for conducting them, from a nationwide perspective in 1998 and 1999 respectively, the Korean government has established and is now operating a new national R&D evaluation system.

### 5.2.2.3. Conclusion

In conclusion, the major events and crucial laws, which have had an important role in the process of the developments in the area of policy evaluation and R&D evaluation since the 1980s, are summarised in <Table 5-3>.

<Table 5-3> Development of Policy Evaluation and R&D Evaluation in Korea

Period Type / Organ.		1980s	1990s	2000s
Policy Evaluation	OPM/ OGPC	- Focused on policy analysis, rather than policy evaluation	- Establishment of PEC ('98) - Evaluation of main policies in central ministries ('98)	- <i>The Framework Act on the Government Performance Evaluation</i> ('01)
	NSTC	-	- Revision of <i>the Special Act for S&amp;T Innovation</i> ('98) - Evaluation of national R&D programmes ('98)	- Evaluation of research councils ('00)
	OPM/ OGPC/ RCs	-	- <i>The Act on the Establishment, Operation and Promotion of GRIs</i> ('98)	- Evaluation of GRIs by RCs ('00) - Evaluation of the S&T related RCs is transferred to OGPC from NSTC ('01)
R&D Evaluation	MOST	- Evaluation of the R&D projects ('82)	- Foundation of STEPI ('92) - Designation of MOST as secretariat of NSTC ('99) - Formation of KISTEP ('99)	- <i>The Framework Act on Science and Technology</i> ('01) - Expansion of KISTEP ('01)

\* *Italics* represent related laws in these sectors

It is apparent that, in the early stage, the national R&D evaluation system in Korea was developed separately from policy evaluation. However, in recent times, the Korean government has tried to make connections between these two evaluation systems. In particular, it appears that the enactment of the Framework Act on Government Performance Evaluation speeded up this trend. As a result, even though lots of improvements remain to be made, from 2001, two evaluation systems started to be incorporated under the responsibility of the OGPC, which is in charge of the evaluation of S&T and R&D policies in the central ministries, in particular the MOST, the MOCIE and the MIC, as well as of the performance of research councils which supervise the majority of Korean GRIs.

### **5.2.3. Current Hierarchy of National R&D Evaluation Systems**

Even though there are many organisations which have carried out evaluation activities in the Korean S&T sector, the NSTC, the OPM/the OGPC, the MOST, the research councils and the KISTEP are the main actors, as noted in Chapter 1. However, each organisation has a different mission in performing evaluation practices. The OPM/the OGPC has two different evaluation responsibilities in the S&T sector. One is to evaluate the S&T and R&D policy of government ministries as a part of the “government performance evaluation” every year (*policy evaluation*),<sup>8</sup> and the other is to conduct the evaluation of the research councils’ performance every year (*institution evaluation*). The NSTC has evaluated all national R&D programmes supported by both the government budget and public funds every year, with administrative assistance from the MOST and the KISTEP (*programme evaluation*). The MOST has also undertaken the evaluation of its own national R&D projects annually, entrusting the actual work to the KISTEP (*project*

---

<sup>8</sup> In Korea, the “progress evaluation” and the “inspection” of the accounts of budgeted projects are conducted by the MPB and the Board of Audit and Inspection (BAI) respectively, separately from the “government performance evaluation”, according to different purposes, standards and procedures.

The MPB is directly and indirectly involved in government policies by drawing up and appropriating government budgets and overseeing budgeted projects. In particular, it is actively involved in the process of formulating government policies by examining the validity of some large-scale budgeted projects from the initiating stage.

Meanwhile, the BAI is a constitutional organisation directly under the President. Its main duties are inspections of accounts and performance of organisational duties. In recent, however, the Board has conducted some policy inspections. The targets of such policy inspections include projects perceived to address social problems and projects for which the President gives special instructions. Such inspections are done on an irregular basis and their main purpose is to realise accountability in administration by determining who is responsible.

For more detail, see Kim (1993) and PEC (2002a).

*evaluation*). On the other hand, the research councils have a principal duty to evaluate the past performance and future research strategies of their member GRIs (*institution evaluation*). At present, it is recognised that the three S&T related research councils have great similarity in their evaluation systems of member research institutes.

The above-mentioned organisations, which conduct different types of evaluation, comprise the upper and lower parts of the national R&D evaluation system, even though the evaluation exercises of different actors now slightly overlap in some components of the evaluation system. Thus, certain evaluation activities such as the institution evaluation of both the OGPC and the research councils should be consistent with the terms and references arising from the evaluation of the main policies in the MOST by the OPM/the OGPC. Also, project evaluation by the MOST should be guided by the results gained from the NSTC's conduct of the evaluation of national R&D programmes. However, at the moment, it seems that there is a lack of close linkage between policy evaluation by the OGPC and R&D programme evaluation by the NSTC, which are considered as the highest level evaluations in each area.

On the other hand, the MPB is in a very special position in national R&D evaluation systems in Korea. Although it is not directly engaged in the actual evaluation work, it plays a crucial role in the process of evaluation utilisation. It is informed of all evaluation findings of four types of evaluation practices, but not of the evaluation of national R&D projects by the MOST. Then, it utilises these findings as one of the policy means to adjust the budget allocation for the coming year and to promote the reforms of the public sector institutions.

The current hierarchy of the national R&D evaluation systems in Korea is shown in (Figure 5-2).



### ***5.2.3.1. Evaluation System for S&T and R&D Policy***

In Korea, the OPM/the OGPC has a principal responsibility for the evaluation of S&T and R&D policies in central ministries as a part of the overall performance evaluation of government ministries. In particular, the evaluation of the performance of the MOST, which co-ordinates all S&T programmes and R&D programmes run by central ministries at the national level, is one of the typical types of evaluation of S&T and R&D policy. In what follows, an overview will be given of the system of the performance evaluation of government ministries in order to give background information regarding S&T and R&D policy evaluation. Then, I will give a more detailed description of “the evaluation of major policy issues” which is a sub area of the evaluation of central administrative offices.

#### **(1) Evaluation of Performance of the Government Ministries by the OPM**

The performance evaluation of the government ministries by the OPM/the OGPC includes the following four main areas : (1) the evaluation of central administrative offices, (2) the evaluation of specific policy tasks, (3) the self-evaluation of the central administrative offices and (4) the evaluation of local governments. For the evaluation of central administrative offices, major policy issues, their capabilities in implementing them and a survey of the public satisfaction with them are evaluated. Regarding the evaluation of specific policy tasks, pressing policy tasks that can emerge at any time are selected and evaluated. The self-evaluation is conducted by both central administrative offices and local governments. Finally, the evaluation of local governments is conducted in relation to affairs delegated by the central government to local governments, projects subsidised from state coffers and other major government policies (OGPC and PEC, 2001a : 1-2 ; PEC, 2002a).

Of the above-mentioned areas, the evaluation of central administrative offices involves the evaluation of main policies implemented by each central ministry. In order to conduct a comprehensive evaluation, this work is again divided into three evaluation areas : (a) major policy issues, (b) institutional capabilities in implementing policies and (c) citizens’ satisfaction with both major policies and civil administrative services. Of the three areas, “the evaluation of major policy issues” directly covers the policy evaluation of each ministry (OGPC and PEC, 2001a : 4-6). Therefore, further explanation will only

be given to the evaluation of major policy issues in central administrative offices.

The evaluation of major policy issues means that core tasks, which are extracted from each ministry's annual plan of duties and are the most crucial and representative of those the ministry performs, are evaluated on a regular basis. For the evaluation, multitasks that are in the process of implementation are normally selected, rather than those in the stage of the policy formulation. Selected programmes and plans should also have practical effects on the public when implemented. In addition, these should be related to the informatisation and knowledge-based society, which are considered as the very foundation of the nation's future competitiveness. In particular, in the first half of the year, evaluations are conducted on policy formulation to determine whether policy objectives and policy means are well linked, on problems arising in the process of implementing policies and on whether countermeasures have been established. In the second half of the year, the focus is placed on the rate of the achievement of the policy objectives and impacts on the public. However, every year, policy issues are adjusted in accordance with the direction of government policy for the year (OGPC and PEC, 2001a : 4-5 ; PEC, 2002).

The process of evaluation comprises the following five stages : (1) the OGPC selects the policy issues to be evaluated, (2) each central ministry sets up implementation plans for its major policies, (3) the OGPC examines the implementation plans for major policy issues and supplements them, (4) the OGPC formulates evaluation plans for each issue, and then conducts the evaluation of each issue twice annually and (5) finally the Prime Minister or the President is briefed on the findings of the evaluation and central ministries are notified of them (OGPC and PEC, 2001a : 5 ; PEC, 2002).

Meanwhile, the evaluation of major policy issues is conducted by reference to six evaluation criteria and 10 evaluation items, depending on the stage in the policy process such as policy formulation, policy implementation and policy achievement. In addition, in the process of the evaluation, use is made of evaluation checkpoints set up in accordance with the policy process. Each ministry must conduct a self-evaluation by using a common evaluation standard before the evaluation exercise undertaken by the OGPC. During this work, if necessary, the ministry can also set up specific extra evaluation items or checkpoints for itself (OGPC and PEC, 2001a : 7).

The findings of the evaluation of government performance are reported in the presentation conference for government evaluation on a six monthly basis. According to the Framework Act on Government Performance Evaluation, they are also open to the public and should be submitted to the National Assembly so that it can refer to them when inspecting the duties of the administration (OGPC and PEC, 2001a : 23).

The findings of the evaluation are intended to be reflected in the development and execution of future policies. First of all, the OGPC notifies the MPB and pertinent ministries about the items that need to be subject to budgetary measures based on the evaluation findings. There is an effort to reflect evaluation findings in the budget process as much as possible. The OGPC also requires pertinent ministries to conduct self-audit of those areas which the evaluation findings indicate are in need of an audit. They are required to report the results of the self-audit. In addition, ministries and staff identified as excellent through the evaluations are provided with incentives, such as awards and preferential treatment in personnel management or budget allocation. On the other hand, unsatisfactory ministries and staff are expected to receive due punishment (OGPC and PEC, 2001a : 23-24).

#### ***5.2.3.2. Evaluation System for R&D Programme***

Regarding the evaluation of R&D programmes at the national level in Korea, the NSTC has the highest level responsibility in the analysis, evaluation and pre-budget review of all national R&D programmes, which have been carried out by central governments since 1998. In addition, some ministries such as the MOST, the MOCIE and the MIC conduct evaluations of their own national R&D programmes. However, this type of evaluation is performed on an irregular basis according to the internal needs of each ministry. So, the following discussion will refer to the evaluation work of the NSTC that is conducted every year.

##### **(1) Evaluation of National R&D Programmes by the NSTC**

This work aims to evaluate the performance and operating structure of national R&D programmes, in order to improve the effectiveness of and to suggest future developments of R&D programmes funded by both national budgets and public funds as a whole. As already noted, even though the NSTC has a formal and legal responsibility for the



evaluation, the MOST, with administrative assistance from the KISTEP, actually conducts this evaluation every year, under the terms of the Framework Act on Science and Technology. In the KISTEP, the Division of National Research and Development Evaluation, one of five internal departments, is in charge of this evaluation work at the operational level. This evaluation was first conducted in 1998, and in the year 2001, the NSTC/the MOST evaluated 161 R&D programmes which had been implemented by 18 central ministries and administrative offices in 2000 (NSTC, 2001a : 94 ; MOST, 2000b ; KISTEP, 2002b).

In principle, the NSTC/the MOST intend to evaluate all R&D programmes supported by the national budgets whether general or special ones as well as public funds including “the informatisation promotion funds” of the MIC and “the nuclear energy R&D funds” of the MOST. Some programmes are carried out merely under the title of research or development programmes, but these are excluded from this evaluation exercise as irrelevant to actual R&D activities. These include the followings : (1) research programmes in the area of the social sciences and humanities, (2) defence-related R&D programmes classified as confidential, (3) programmes for R&D planning and management, R&D policy studies and field surveys in each ministry (4) programmes for their organisational remits in R&D management institutes such as the KISTEP, the ITEP and the IITA and (5) programmes for construction and refurbishment of research buildings (MOST, 2000b ; KISTEP, 2002b). As a result, through this evaluation practice, the Korean government tries to evaluate R&D programmes in the public sector to the greatest extent possible. All R&D programmes that are selected for evaluation are grouped into four main programme groups and 12 sub programme groups, depending on their nature and characteristics (MOST, 2000b : 5-6 ; KISTEP, 2002b).

For this evaluation, “Evaluation Committees for sub R&D Programmes” are set up according to sub programme groups. Each committee, comprising around 10 external experts from academia, research institutes and industry, has three missions : to evaluate R&D programmes in each group, to conduct a final review of linkages and overlapping among R&D projects and to co-ordinate priorities among R&D programmes falling within its remit. They conduct an evaluation on the basis of both the presentation made by a programme manager and reports submitted by the related ministry. They also adopt three evaluation criteria, namely the appropriateness of the programme objectives, the

efficiency of implementation and the level of performance and outputs. Each criterion is divided into three to six sub criteria. For each criterion, an individual grade is given according to a five-tier grading system and the reason for that grade is also provided as well as recommendations for future improvements in the qualitative evaluation method. Finally, the committee gives each individual programme a grade on five-tier grading system so that it can be compared with other programmes within the same programme group (MOST, 2000b : 11 ; KISTEP, 2002b).

As already noted, this work comprises a major activity in the cycle of the analysis, the evaluation and pre-budget review of national R&D programmes. Therefore, the findings of this evaluation are closely linked with the pre-budget review of the pertinent R&D programmes for the coming year. That is, R&D programmes that have made excellent progress are favourably treated in the process of budget allocation for the coming year. Also, those who have made a great contribution to the pertinent R&D programmes are given an award. In addition, through the synthesis of findings of the analysis and evaluation of national R&D programmes, the NSTC presents the science and research areas that require government investment and suggests funding priorities among R&D programmes (MOST, 2000b : 3 ; KISTEP, 2002a).

#### ***5.2.3.3. Evaluation System for Research Institutions***

There are two types of practice in the evaluation of institution in the science and technology sector in Korea. These have similar evaluation methods, evaluation criteria and so on, even though the evaluation objects are different from each other. One is the “evaluation of the five research councils” in which the OPM/the OGPC evaluates the performance of its own businesses and its mission of supporting the GRIs managed by them, and reviews the appropriateness of the internal management system every year. The other is an “evaluation of the research institutes, the GRIs” in which all research councils evaluate both the research performance of the GRIs including research quality and the rationality of their future R&D strategies every year.

In the case of the evaluation of research institutes controlled by research councils, the Korean practice is different from that of the UK. While the UK research councils have their own differentiated evaluation systems depending on the characteristics of their

supporting research communities, the Korean research councils all have a common evaluation system. For this reason, in contrast to the UK case in which the Institute Assessment Exercise (IAE) by the BBSRC is selected as a case study, I will not choose a specific example for the Korean case study.

### **(1) Evaluation of Research Councils by the OPM/the OGPC**

The OPM/the OGPC has evaluated the performance of five research councils' own businesses and the mission of supporting the GRIs managed by them and has reviewed the appropriateness of internal management system every year since 2000. This evaluation is conducted to ensure that the output of the research councils' businesses is maximised and that the effectiveness of their management system is enhanced. For this work, a common evaluation manual, which is applicable to all research councils and is open to them at the beginning of the year, has been adopted (OPM, 2001c ; OGPC, 2001a : 1-2).

To do this work, the "Committee for the Evaluation of the Research Councils", which is chaired by a private sector expert and consists of around 10 experts from academia, research institutes, the media and industry, is established temporarily. It carries out the evaluation of five research councils with assistance from an external institution that has professionalism in the area of R&D management, including research evaluation, and acts as the secretariat on the basis of a contract. For the evaluation, the committee normally reviews the report submitted by each research council and visits each council as a complementary evaluation method (OGPC, 2001a : 4).

The evaluation objects for the research councils are categorised into three groups : (1) research councils' businesses and the mission of supporting their member GRIs, (2) internal management systems and (3) a satisfaction by related institutions. Each evaluation object is given a different weight according to its importance in the perspective of the evaluation purposes. That is, 70% of the weight falls on the research councils' businesses and the mission of supporting their GRIs, whereas 30% falls on the internal management system. Again, three evaluation objects are divided into 10 sub evaluation objects, reflecting the nature and contents to be evaluated. The evaluation objects comprise 21 evaluation indicators and 36 evaluation items that are described in terms of an operational definition for actual evaluation work. For each evaluation item, a score is

given on a five grade scale by using a qualitative method mixed with a quantitative one (OGPC, 2001a : 3, 5-7).

The OPM/the OGPC publishes an evaluation report and informs the MPB of evaluation findings. In particular, those in relation to the three research councils in science and technology sector, namely the KRCF, the KORP and the KOCI, are notified to the NSTC as well as to the MPB (OPM, 2001c ; OGPC, 2001a : 1). Through this process, the Korean government utilises evaluation findings as a policy tool to adjust the budget for public sector R&D activities as well as to co-ordinate the national R&D programmes.

## **(2) Evaluation of Research Institutes (GRIs) by the Research Councils**

Five Korean research councils have conducted the evaluation of their member GRIs every year since 2000. Through this evaluation, they intend to provide member institutes with a forward direction for their developments, to induce them to reform their internal management system actively in order to improve research productivity, to enhance their capabilities continuously and actively in order to create value and to disseminate their research outputs and knowledge. In order to carry out the evaluation of the GRIs as a collective task at the national level, in particular, the three S&T related research councils prepare a common evaluation manual in collaboration with the OGPC. Meanwhile, to ensure the credibility and expertise of an evaluation, this work is normally commissioned to an external institution which specialises in the area of R&D evaluation (OPM, 2001c ; KOCI, 2002b : 1, 4).

This evaluation is also conducted by “the Committee for Evaluation of Research Institutes” as like the evaluation of research councils. The committee is chaired by a external expert and comprises around 15 professionals from academia, research institutes, the consulting company and industry. It is again divided into two sub-committees, one mainly responsible for R&D matters and the other for internal management matters. They are also supported by an external institution to which work is commissioned by research councils on the basis of a contract. In accordance with the evaluation method, the committee reviews the performance report submitted by each institute and visits each GRI in order to confirm the contents of its report and to conduct some interviews with related staff (KOCI, 2002b : 6).

The evaluation areas for research institutes are categorised into two groups : R&D matters and internal management matters. In contrast to what occurs in the evaluation of research councils, each evaluation area is given the same weight. They are again divided into various evaluation criteria and 22 evaluation indicators, which are given a different weight. While all research councils adopt the same structure and terminology in relation to evaluation indicators, they also have autonomy to decide the weight of each indicator differently, depending on its nature and the content to be evaluated (KOCI, 2002b : 10-13).

Each research council publishes its own evaluation report and submits evaluation findings to both the Prime Minister and the minister of the MPB. These findings can be utilised, where appropriate, to close any institute that shows bad progress in the evaluation work as well as to lay off the president of an institute who is considerably lacking in management ability and the capability necessary for his position. In addition, each institute has to reflect the evaluation findings in the strategies and plans of the institute management and research programmes for the following year (OPM, 2001c ; KOCI, 2002b : 6).

#### ***5.2.3.4. Evaluation System for National R&D Project***

As already noted, in Korea, some S&T related ministries including the MOST, the MOCIE and the MIC, have carried out national R&D programmes for their own policy purposes. In order to manage these programmes effectively, they have also established R&D management institutes as affiliated organisations such as the KISTEP, the ITEP and the IITA. As a result, these ministries conduct the evaluation of R&D programmes by entrusting their evaluation missions to R&D management institutes. Of these, the MOST has a longer history and a well-established system of evaluation, which started with the initiation of the first Korean national R&D programme in 1982. The experiences of the MOST in this field have also influenced other ministries to set up their own evaluation systems and to improve evaluation practices.

Meanwhile, the actual evaluation practice in each ministry was not targeted at the R&D programme itself but at R&D projects, a sub component of the programme. In Korea, at ministry level, the evaluation of R&D projects is conducted on a regular basis, whereas

the evaluation of the R&D programme is conducted irregularly. That is, the related ministries conduct an *interim* evaluation at the end of each funding year, as well as an *ex-post* one after the completion of a pertinent R&D project.

For these reasons, discussion will now focus on the evaluation system of the MOST and, in particular, on that of R&D projects.

### **(1) Evaluation of National R&D Projects by the MOST**

The national R&D programme of the MOST, initiated in 1982, has its legal basis in both “the Technology Development Promotion Act” and “the Framework Act on Science and Technology”. Through the 1980s and 1990s, the MOST developed diverse national R&D sub-programmes in order to meet socio-economic needs in an informatisation and knowledge-based society. Of these, the Highly Advanced National Project (The HAN Project), which was launched in 1992 as an inter-ministerial programme and funded by both government and industry, has been considered a flagship programme. The national R&D programme supported by the MOST currently comprises some specific sub-programmes including (1) The 21st Century Frontier R&D Programme, (2) The HAN Project, (3) The Creative Research Initiatives (CRI), (4) The National Research Laboratory (NRL) Programme, (5) The Biotechnology Development Programme, (6) The Space and Aeronautics Programme and so on (MOST, 2002c).

The MOST has a principal policy to conduct an “appraisal” in order to select R&D projects, to “monitor” a project during its implementation, to undertake an “output evaluation” just after its completion and an “impact evaluation” some years after its completion. Therefore, three types of R&D evaluation are conducted in relation to R&D projects funded by the MOST. In particular, the MOST carries out an interim evaluation, namely monitoring, of R&D projects which are supposed to be supported for over two years. Meanwhile, according to “the Technology Development Promotion Act” and “the Regulations on Management of National R&D Programme by the MOST”, the evaluation mission at the operational level is entrusted to the KISTEP. In the KISTEP, the Division of National R&D Planning and Management, also one of the five departments, is responsible for this work (MOST, 2002f ; MOST 2002g).

Like other Korean evaluation system, this evaluation is conducted by at least seven

external experts from academia, research institutes and industry. They usually conduct an evaluation on the basis of a final research report and a self-evaluation report submitted by a pertinent institute, whilst the project leader gives an open presentation of research outputs. Then, on the basis of evaluation results of external experts, the KISTEP gives a certain grade on a five-tier grading system (MOST, 2002g).

Regarding the utilisation of evaluation findings, firstly, those R&D projects that had a bad evaluation, if continue to be supported, may fail to get further funds for the coming year. Next, on the basis of the evaluation findings, it is expected that the KISTEP may have to make a decision about whether the project leader will be given an incentive or a punishment in the future. In addition, research institutes that carried out pertinent R&D projects have to reflect these in the internal evaluation practices of researchers involved in those projects. Meanwhile, the Minister of the MOST is assigned to publish a “white paper on evaluation” which contains some important matters of evaluation process and exercise (MOST, 2002f ; MOST, 2002g).

#### ***5.2.4. Current Referential Statutes for R&D Evaluation***

In contrast to the UK, in Korea, all government administrations have to be executed on the basis of statutes such as laws, presidential decrees, ministerial ordinances and detailed enforcement regulations. Hence, evaluation systems have been designed and actual evaluation exercises have also been conducted under the related statutes.

The central government organisations, including the OPM/the OGPC, the MOST and the MPB, have enacted referential statutes and, have sometimes issued basic guidance on the execution of both policy evaluation and R&D evaluation. Along with the central organisations, individual ministry, for example, the MOST, has also enacted enforcement regulations at the operational level, which deal with more specific and detailed matters delegated to the ministry by higher level statutes. The current referential statutes for R&D evaluation, including the evaluation of S&T and R&D policies, are summarised in <Table 5-4>. In this part, I shall give a brief explanation of the current statutes at the level of both central governments and executive organisations.

**<Table 5-4> Current Referential Statutes for R&D Evaluation in Korea**

<b>The Evaluation of S&amp;T and R&amp;D Policies</b>
<ul style="list-style-type: none"> <li>- The Framework Act on Government Performance Evaluation and its Enforcement Decree (2001)</li> <li>- The Regulations on Operation of the Policy Evaluation Committee (2001)</li> </ul>
<b>The Evaluation of National R&amp;D Programmes</b>
<ul style="list-style-type: none"> <li>- The Framework Act on Science and Technology, its Enforcement Decree and Ministerial Ordinance (2001)</li> <li>- The Regulations on Operation of the Evaluation Committee for National R&amp;D Programmes (2001)</li> </ul>
<b>The Evaluation of Research Councils and Research Institutes</b>
<ul style="list-style-type: none"> <li>- The Act on Establishment, Operation and Promotion of the Government-supported Research Institutes (GRIs) and its Enforcement Decree (2001)</li> </ul>
<b>The Evaluation of National R&amp;D Projects by the MOST</b>
<ul style="list-style-type: none"> <li>- The Framework Act on Science and Technology, its Enforcement Decree and Ministerial Ordinance (2001)</li> <li>- The Presidential Decree on Planning, Management and Evaluation of National R&amp;D Programmes (2001)</li> <li>- The Regulations on Management of National R&amp;D Programme by the MOST (2002)</li> <li>- The Regulations on Evaluation of National R&amp;D Projects by the MOST (2002)</li> </ul>

While some statutes cover only one evaluation system, others concern the affairs of two evaluation systems. The former is the Framework Act on Government Performance Evaluation which is the reference for policy evaluation. The latter include the Framework Act on Science and Technology which is the reference for the evaluation of national R&D programmes and projects, and the Act on Establishment, Operation and Promotion of the Government-supported Research Institutes (GRIs) which is the reference for the evaluation of research councils and research institutes. Even though some statutes relate to more than two evaluation systems, a description will be given to each individual evaluation system separately.

#### ***5.2.4.1. The Statutes for the Evaluation of S&T and R&D Policies***

Regarding the evaluation of S&T and R&D policies, three types of statute are now in operation. These include “the Framework Act on Government Performance Evaluation”, its “Enforcement Decree” and “the Regulations on Operation of the Policy Evaluation



Committee". Among these, the Framework Act on Government Performance Evaluation relates to the basic policy and some important principles in relation to the evaluation of government performance including S&T and R&D policies, whereas its Enforcement Decree and the Regulations on Operation of the Policy Committee deal with more specific matters at the acting level that are delegated to the President and the Minister of the OGPC for the more effective execution of the evaluation.

**(1) The Framework Act on Government Performance Evaluation and its Enforcement Decree (OPM, 2001)**

As already noted, to establish an efficient evaluation system for government policies and programmes, and to promote the efficiency and integrity of evaluation activities in local governments through general evaluation standards, the Korean government enacted the Framework Act on Government Performance Evaluation in January 1999 and its Enforcement Decree in May 2001.

The main contents of this law and its presidential decree are summarised as follows (OPM, 2001a ; PEC, 2002a) :

- (1) Evaluation Principles : The law was legislated to guarantee the autonomy, independence and objectivity of the government's evaluation exercises and to prevent repeated and burdensome evaluations by multiple evaluation organisations.
- (2) Evaluation Area : There are four main areas of evaluation described in this law : the evaluation of central administrative offices, the evaluation of specific policy tasks, self-evaluation by both central and local government and the evaluation of local governments. (a) For the evaluation of central administrative offices, the major policy issues of each ministry and agency and its capabilities are evaluated. In addition, the results of a survey of public satisfaction are evaluated. (b) Regarding the evaluation of specific policy tasks, pressing policy tasks that emerge at any time are selected and evaluated. They include policy tasks arising in response to social change such as the shift to an information and knowledge society and cases that need close examination for effectiveness because of large budgets and so on. (c) Self-evaluations are conducted by both central administrative offices and local governments. That is, the government, including local governments, conducts an independent evaluation of its capabilities in promoting policies related to their main functions and in implementing an innovative style of management. This is aimed at establishing a system to identify

problems in advance and to stress them internally. (d) The evaluation of local governments is conducted on the affairs delegated to it by central government, projects subsidised from state coffers and other major government policies. If necessary, the Minister of Government Administration and Home Affairs, which is in charge of the general affairs of local governments, may conduct a joint evaluation in conjunction with related central administrative offices.

- (3) Evaluation Procedures : Firstly, the Office for Government Policy Co-ordination (OGPC) establishes annual evaluation guidelines by 31 January each year. Next, the heads of central government ministries and agencies submit their evaluation plans to the OGPC by 10 March. According to these guidelines and plans, the Policy Evaluation Committee holds meetings to conduct the evaluations and make their evaluation report.
- (4) Evaluation Utilisation : In accordance with the evaluation findings, the heads of the ministries and agencies are asked to rectify problems and conduct inspections. The Minister of the MPB, which is responsible for government budget allocation, is notified about the items requiring budget adjustments. The OGPC also follows up in the following year on the items that were given low evaluation results in the previous year.
- (5) Disclosure of Evaluation Findings : In principle, to enhance public surveillance over policies, those who conduct the evaluations are required to disclose the findings. However, some institutions covered by the non-disclosure articles of the Information Disclosure Act are excluded from this principle of disclosure.
- (6) Policy Evaluation Committee : Under the Prime Minister, there is the Policy Evaluation Committee (PEC) which comprises prominent private sector experts and is charged with carrying out the evaluations of government performance.
- (7) Government Performance Evaluation Consultative Meeting : This meeting, chaired by the Minister of the OGPC, consults on evaluations and co-ordinates ways to efficiently promote evaluations among major government ministries and agencies.

## **(2) The Regulations on Operation of the Policy Evaluation Committee (OPM, 2001)**

As already noted, the Korean government established the Policy Evaluation Committee under the Prime Minister in 1998, in order to secure the objectivity and professionalism of evaluation work through the involvement of experts from the private sector and to recommend developments for the evaluation system. The remit of the committee is to

review and decide on the following matters : (1) the basic directions and guidelines for government performance evaluation, (2) the conduct of evaluations of central ministries and administrative offices, local governments and specific policy tasks and (3) actual management and development of the evaluation system (PEC, 2002a ; OPM, 2001b).

The committee is chaired by an expert from the private sector and comprises around 30 members from academia, research institute, industry and the media. From the government, only the Minister of the OGPC sits on this committee. The commitment period of committee member is normally two years. In order to deal with the administrative affairs of the committee, the Assistant Minister for Policy Analysis and Evaluation, one of five Assistant Ministers under the OGPC, acts as the secretariat. Meanwhile, the committee is divided into six sub-committees, which assist the main committee by handling more detailed policy issues. These include (1) the sub Committee for Evaluation Systems and Management, (2) the sub Committee for Economic Policies I, (3) the sub Committee for Economic Policies II, (4) the sub Committee for Social and Cultural Policies, (5) the sub Committee for General Administrative Policies and (6) the sub Committee for Local Government. In addition, the committee may set up working groups to conduct the evaluation of specific policies (OPM, 2001b).

#### ***5.2.4.2. The Statutes for the Evaluation of National R&D Programmes***

Regarding the evaluation of national R&D programmes, four kinds of statute are now in force. Of these, “the Framework Act on Science and Technology”, its “Enforcement Decree” and “Ministerial Ordinance” contain basic contents for the evaluation of national R&D programmes at the national level, while “the Regulations on Operation of the Evaluation Committee for National R&D Programmes” deals with more detailed and specific matters in relation to the evaluation committee which have been delegated by the above mentioned statutes to the Minister of the MOST.

##### **(1) The Framework Act on Science and Technology, its Enforcement Decree and Ministerial Ordinance (MOST, 2001)**

This act is the basic law covering the systematic promotion of S&T innovation at national level. Important provisions of the act include the establishment of policies and plans for S&T and overall support mechanism for related projects and agencies. The act also aims

to provide the legal reference for inter-ministerial coordination of R&D activities and to establish an institutional system to foster an innovation-friendly culture in Korean society. This act supersedes and includes the major contents of two previous S&T related laws, “the Science and Technology Promotion Act” and “the Special Act on Scientific and Technological Innovation”.

Some of the provisions of these statutes in relation to the evaluation of national R&D programmes will now be briefly described (MOST, 2001a).

- (1) Planning and Execution of National R&D Programme : The S&T related central ministries have a responsibility for performing national R&D programmes in their policy areas. When the government plans and performs a national R&D programme, it must consider the needs of related industries. It must also set up common principles and standards for the planning, evaluation and management of the national R&D programmes, in order to implement the programmes transparently, fairly and effectively and to create close links among national R&D programmes supported by individual central ministries.
- (2) Operation of the National Science and Technology Council (NSTC) : The government runs the NSTC in order to review and co-ordinate national S&T policies and R&D programmes and to set priorities for the allocation of R&D budgets at national level. As already noted in Section 5.2, the NSTC is chaired by the President and comprises at most 24 members who come from S&T-related ministers as well as from the representatives of S&T community including industry, academia and research institutes. Of these members, the Minister of the MOST acts as the executive secretary of the NSTC. At the same time, in order to deal with specific S&T related matters, the government sets up the “Operating Committee” and the “Special Committee” under the NSTC. While various expert committees, for example, the Expert Committee for S&T policy, the Expert Committee for R&D and the Expert Committee for Aerospace Development are answerable to the operating committee, the Committee of the Bio-technology and Bio-industry is answerable to the special committee.
- (3) Utilisation of the Review Findings of the NSTC : The NSTC should communicate its review findings to related central government ministers and governors of local governments. They are also asked to reflect review findings in S&T policy and in the allocation of S&T related budgets under their control.

- (4) Analysis and Evaluation of National R&D Programmes : As the secretariat of the NSTC, in principle, the Minister of the MOST has to conduct the analysis and evaluation of all national R&D programmes and to report its findings to the NSTC. For this work, he/she can ask central ministries, local governments, government-invested institutions, higher education institutions, research institutes and private companies involved in national R&D programmes to submit necessary materials and information. He/she also sets up the plan for the analysis, evaluation and budget pre-view of national R&D programmes every year. Then, the minister must report this plan to the NSTC and notify it to related ministers and the chairpersons of the three S&T related research councils by 30th November every year. The related ministers in central government should reflect findings of this work in the execution of R&D programmes, in order to ensure the effectiveness of R&D investment. Meanwhile, the NSTC has to review both the plan for national R&D programmes and the demands for R&D budget for the coming year taking into account the findings of the analysis and evaluation of national R&D programmes in the previous year. Then, the Council must notify these results to the Prime Minister, related ministers in central government and the Minister of the MPB.
- (5) Establishment of the Korea Institute of S&T Planning and Evaluation (KISTEP) : The KISTEP was established, in order to support the analysis, evaluation and pre-budget review of national R&D programmes conducted by the MOST on behalf of the NSTC. In addition to its mission in relation to the NSTC, the KISTEP is also responsible for supporting management of national R&D programme by the MOST.
- (6) Promotion and Evaluation of the Government-supported Research Institutes (GRIs) : The government should endeavour to promote the GRIs, in order to carry out national R&D programmes effectively. On the other hand, the Prime Minister and related ministers conduct the evaluation of the GRIs supported by both the research councils and other central ministries and submit the findings of this evaluation to the NSTC.

**(2) The Regulations on Operation of the Evaluation Committee for National R&D Programmes (MOST, 2001)**

This regulations aims to deal with detailed matters regarding operation of the evaluation committee that has a responsibility for conducting analysis and evaluation of national R&D programmes based on the Framework Act on Science and Technology. According to the regulations, there are three types of sub committee under the evaluation committee

for national R&D programmes. These are a “Steering Committee” to co-ordinate and integrate evaluation findings, a “Review Committee for R&D Projects” to review the linkage and overlapping among R&D projects and an “Evaluation Committee for sub R&D Programmes” to conduct the evaluation of sub R&D programmes (MOST, 2001c).

- (1) **Steering Committee** : This committee reviews and co-ordinates results from the review committee of R&D projects and the evaluation committee of sub R&D programmes. It is chaired by a private sector member from the NSTC and consists of at least 20 members including private sector members of the operating committee, chairpersons of the expert committee for S&T policy and the expert committee for R&D and chairpersons of the evaluation committees of sub R&D programmes.
- (2) **Review Committee for R&D Projects** : The mission of this committee is to review the possibility of linkage and overlapping among R&D projects. The committee is organised to cover around 17 main S&T areas, namely basic science, biotechnology, IT technology and so on. It is chaired by a private sector expert and comprises around 10 experts from industry, academia and the research institutes.
- (3) **Evaluation Committee for sub R&D Programmes** : This committee conducts the evaluation of sub R&D programmes and co-ordinates the results from the review committee of R&D projects. The committee is organised to handle major sub R&D programmes under entire national R&D programmes, such as ① basic, publicly-used and welfare technology programmes, ② industrial technology R&D programmes, ③ R&D infrastructure building programmes and ④ research institutes support programmes. It is also chaired by a private sector expert and comprises around 10 experts from industry, academia and the research institutes.
- (4) **Operation of Supporting Team** : In order to provide several committees with administrative supports and to ensure the fairness of the analysis and evaluation of national R&D programmes, the government may set up “Supporting Team”. This team comprises around 15 R&D management experts and is headed by the managing director of the “Division of National Research & Development Evaluation” in the KISTEP.
- (5) **Secretariat of the Committee** : The KISTEP, under the Framework Act on Science and Technology, is designated as the secretariat of the evaluation committee.

#### ***5.2.4.3. The Statutes for the Evaluation of Research Councils and GRIs***

The terms of reference of the evaluation of research councils and their member research institutes can be extracted from the Act on Establishment, Operation and Promotion of the GRIs and its Enforcement Decree. In contrast to other cases, there are no specific regulations in relation to these works. However, these statutes relate only to the evaluation of research councils and the GRIs supervised by them. As a result, the evaluation of the GRIs under other ministries' responsibility, including the MOST, is conducted on the basis of the Framework Act on Science and Technology.<sup>9</sup>

##### **(1) The Act on Establishment, Operation and Promotion of the Government-supported Research Institutes (GRIs) and its Enforcement Decree (OPM, 2001)**

This law and presidential decree deal with the establishment, promotion and accountability of the GRIs and provide a systematic management system for them so as to facilitate them in their aims, to promote the rationalisation and development of the GRIs' management system and to secure the rationality of the national R&D system. To carry out these missions, the government adopts the evaluation of research councils and their research institutes as a part of its diverse policy means (OPM, 2001c).

- (1) Evaluation of Research Councils : The Prime Minister conducts the evaluation of three S&T related research councils every year and notifies both the NSTC and the Minister of the MPB of the evaluation findings. The main evaluation object consists of two types of research council activities. One is the supervision and management of their member institutes, while the other concerns their organisational missions which include R&D planning in S&T areas with their responsibility, a plan for future the development of member institutes, coordination of organisational functions among member institutes, evaluation of research outputs and management performance, support for collaborative research among member research institutes and so on.
- (2) Evaluation of GRIs : Each research council should conduct the evaluation of research outputs and management performance of member research institutes every year. To ensure the objectivity and fairness of the evaluation, the research council may commit this work to an internal "Planning and Evaluation Committee" or an external

---

<sup>9</sup> As already noted, there are two types of the GRIs in Korea. Firstly there are the research institutes under the supervision of the five research councils. Secondly there are the research institutes supported by central ministries including the MOST, the MOCIE, the MOD and so on.

institution with high expertise in the area of R&D evaluation. The act and presidential decree provide very detailed categories of evaluation objects. These include ① the selection of R&D projects, ② the execution and management system for R&D projects, ③ a specialisation in research fields, ④ a collaborative research between industry and academia, ⑤ the quality of research outputs, ⑥ the exploitation and dissemination of research outputs, ⑦ the setting-up of management objectives and their achievement, ⑧ the management of internal organisation and human resources, ⑨ the management of the finance and R&D budget and ⑩ the rationalisation of management system in order to improve R&D productivity. Then, the research councils are expected to submit evaluation findings to the Prime Minister and the Minister of the MPB.

- (3) Planning and Evaluation Committee : The research councils may set up a “Planning and Evaluation Committee”. Such a committee comprises around 10 experts from industry, academia and the research institutes. It coordinates missions among member institutes, provides them with advisory comments in relation to the long-term development visions and supports evaluation practices of them.

#### ***5.2.4.4. The Statutes for the Evaluation of National R&D Projects***

There are two types of statutes in relation to the evaluation of national R&D projects. There are firstly higher-level statutes that cover all R&D evaluation systems of related ministries at the national level, and secondly lower-level regulations that deal with more concrete matters delegated to individual ministers by higher-level statutes. The former include the “Framework Act on Science and Technology,” its “Enforcement Decree” and “Ministerial Ordinance” as well as the “Presidential Decree on Planning, Management and Evaluation of National R&D Programme”. In accordance with higher-level statutes, some ministries issue detailed regulations for the effective evaluation of national R&D projects. In particular, the MOST, the most important ministry responsible for S&T and R&D affairs, has enacted specific regulations including “the Regulations on Management of National R&D Programme” and “the Regulations on Evaluation for National R&D Projects”.



**(1) The Framework Act on Science and Technology, its Enforcement Decree and Ministerial Ordinance (MOST, 2001)**

As already noted, these statutes deal with the evaluation of both national R&D programmes at the national level and national R&D projects at individual ministries level. As an account has already been given of the evaluation of national R&D programmes, further discussion will be confined to the evaluation of national R&D projects of each ministry level.

The following two provisions in these statutes deal directly with the evaluation of national R&D projects (MOST, 2001a).

- (1) Planning and Execution of National R&D Programme : The S&T related central ministries have a responsibility for performing national R&D programmes in their policy areas. The government is also expected to set up common principles and standards for the planning, evaluation and management of national R&D programmes, in order to carry out the programmes transparently, fairly and effectively and to make close linkages among national R&D programmes supported by individual central ministries. According to these provisions, some central ministries, for example, the MOST, the MOCIE, the MIC and the MOHW, carry out diverse national R&D programmes under their accountability and run evaluation systems of national R&D projects when they consider them necessary.
- (2) Establishment of the Korea Institute of S&T Planning and Evaluation (KISTEP) : The KISTEP, established under this act, has two missions which are slightly different, but closely connected. One is to act as the secretariat for the NSTC, the other is to manage the MOST's national R&D programme which comprises sub R&D programmes including the 21st Century Frontier R&D Programme, the HAN Project, the Creative Research Initiatives (CRI), the National Research Laboratory (NRL) Programme and so on. Therefore, it has a responsibility for the evaluation of national R&D sub programmes and R&D projects funded by the MOST.

**(2) The Presidential Decree on Planning, Management and Evaluation of National R&D Programmes (MOST, 2001)**

This decree has been enacted to deal with the planning, management and evaluation of individual ministries' national R&D programmes, which have been delegated to the President by Article 11 of the Framework Act on Science and Technology. According to

this decree, a national R&D programme can be defined as the R&D programme and projects in the S&T sector both fully or partially subsidised from state coffers and funded by public funds including the informatisation promotion funds of the MIC. The decree contains some crucial matters regarding management stages of the national R&D programme, including research planning, project selection, the submission of final outputs, evaluation and its utilisation, the exploitation and knowledge transfer of research results and so on. Of these matters, further explanation will be given regarding only the monitoring and evaluation of R&D project, not their appraisal.

The central government minister has to conduct an “interim evaluation”, effectively a monitoring exercise, and an “ex-post evaluation” of national R&D projects. In addition, in order to trace the exploitation and knowledge transfer of the research results, he/she may conduct an “impact evaluation” some years after projects completion. However, some R&D projects incorporating matters of national security may be excluded from the evaluation. In conducting interim and ex-post evaluation, in principle, the principal investigator should make an open presentation of the research outputs and the evaluator should adopt an appropriate evaluation method, which can be used to judge the output of a pertinent project by comparison with others and to give them a ranking. Also, the central ministries will make an effort to ensure the professionalism, objectivity and fairness of the evaluation process and its findings, for example, by adopting international experts as evaluators, where appropriate.

The minister may stop research funds of R&D projects that have been shown by the evaluation findings to have made bad progress. He/she can also punish the principal investigator of the pertinent project, who will not be allowed to carry out another national R&D programme for two years. On the other hand, he/she may give excellent evaluation projects some incentives including support for the exploitation of research outputs. In addition, the minister may regularly publish a “white paper on R&D evaluation”. Meanwhile, the ministry should set up an appeal system, which allows investigators who are dissatisfied with the results to complain about the outcomes of the evaluation or appraisal (MOST, 2001b).

### **(3) The Regulations on Management of National R&D Programme by the MOST (MOST, 2002)**

This regulations aims to lay down detailed contents necessary for the effective management of the national R&D projects of the MOST, which are based on the Framework Act on Science and Technology, the Technology Development Promotion Act and the Presidential Decree on Planning, Management and Evaluation of National R&D Programme. Therefore, the regulations covers overall R&D management matters including annual research planning, the selection of R&D projects, the management of research funds, the evaluation of research outputs, the exploitation and knowledge transfer of research outputs and so on. However, the following account will concentrate on evaluation including monitoring.

In principle, in order to decide whether R&D projects that are supposed to be supported for over two years will continue to be funded or not, the MOST monitors or conducts an output evaluation of pertinent projects. In fact, under “the Technology Development Promotion Act” and this regulations, the evaluation function is entrusted to the KISTEP. On the basis of the evaluation findings, R&D projects that are evaluated as being poorly conducted may fail to get further funds for the coming year (MOST, 2002f).

### **(4) The Regulations on Evaluation for National R&D Projects by the MOST (MOST, 2002)**

This regulations was enacted to handle the appraisal, monitoring, output evaluation and impact evaluation of R&D projects funded by the MOST, in order to support the effective implementation of national R&D projects. It contains very detailed provisions for overall R&D assessments.

As in the case of other statutes, an account will now be given of evaluation work including monitoring only (MOST, 2002g).

- (1) Principles of Evaluation : Through the introduction of this regulations, the MOST intends to secure some important principles such as enhancing the generality, expertise and fairness of evaluation work, strengthening the R&D management system by conducting strict evaluation, establishing an open evaluation system on which researchers may suggest their opinions during the evaluation process and managing a differentiated evaluation system to reflect characteristics of individual

R&D projects.

- (2) Interim Evaluation (Monitoring) : The KISTEP conducts a monitoring or an interim evaluation of annual outputs and future plans during the performance of R&D projects. It may select evaluators either from members of an exclusive evaluator group or from experts from industry, academia and research institutes. For the monitoring, the research institute responsible for pertinent R&D projects makes an open presentation, where evaluators attend to conduct an interim evaluation. In principle, the KISTEP refuses further research funds to the few R&D projects, which showed the worst evaluation in monitoring and, where appropriate, site visits.
- (3) Output Evaluation : Output evaluation is conducted by at least seven external experts from academia, research institute and industry. In the selection of the evaluators, in principle, members of “an exclusive evaluator group” or experts engaged in previous appraisal and monitoring stages are preferred. If necessary, the KISTEP may also select evaluators from other professionals or, where appropriate, from international experts. They usually conduct an evaluation on the basis of a final research report and a self-evaluation report submitted by an institute, whilst the project leader gives an open presentation about research outputs. Then, on the basis of the results of the evaluation by external experts, the KISTEP gives a certain grade on a five-tier grading system and makes a decision about whether to give incentives or impose a punishment.

Regarding the evaluation utilisation, firstly, if it is supposed that R&D projects that have shown bad progress are likely to continue to do so, they may fail to get further funds for the coming year. Next, on the basis of the evaluation findings, it is expected that the MOST and the KISTEP may have to take a decision about whether the project leader will be given an incentive or a punishment, when he/she will intend to apply for funds for other R&D projects in the future. In addition, research institutes that have carried out pertinent R&D projects will be asked to reflect these in the internal evaluation exercises of the researchers involved in them.

- (4) Impact Evaluation : In order to encourage the exploitation and knowledge transfer of research outputs, the MOST may conduct an impact evaluation three years after the completion of R&D projects. This evaluation will normally be conducted by three experts including a business expert, a related member of staff of the KISTEP and an external technical expert. R&D projects with an excellent evaluation will be given the incentives necessary for the commercialisation of the research outputs and so on.

- (5) Evaluation of National R&D Programme : The MOST may carry out the evaluation of the national R&D programme itself, in order to connect upper level programmes with lower level ones as well as to judge research performance and whether the R&D programme have achieved its goal. In principle, this evaluation will be conducted every three years, even though this varies depending on the research planning set up in advance.
- (6) White Paper on Evaluation Work : The Minister of the MOST is expected to publish “a white paper on evaluation” which contains an evaluation plan, lists of projects evaluated, evaluation procedures, evaluation findings, a list of evaluators and so on every two or three years.
- (7) Establishment of Exclusive Evaluator Group : In order to improve the continuity and expertise of R&D management, the KISTEP may set up “an exclusive evaluator group” which comprises around 10 experts and will be engaged in all stages of an appraisal, a monitoring and an output evaluation. In principle, such a group must be set up before R&D project is selected, but exceptionally it can also be established during the performance of projects if this is found necessary.

### **5. 3. Recent Evaluation Exercises in S&T and R&D Sectors**

In Korea, according to some referential statutes of the central governments as well as the executive ministries, diverse evaluation activities have been conducted at different levels of the S&T and R&D sectors. As already noted, these evaluations are grouped into four types : “the evaluation of S&T and R&D policy” as a part of government performance evaluation, “the evaluation of national R&D programmes” at the national level, “the evaluation of institution” (e.g. research councils and their member research institutes, namely GRIs) and “the evaluation of national R&D projects” at individual ministry level. Consequently, in Korea, five different types of evaluation have normally been carried out in the S&T and R&D sectors.

As already noted, the MOST has a well-established evaluation system concerning the R&D project. Among its R&D programmes, the CRI programme is considered to have an advanced evaluation system that is very different from that of other programmes. It introduced, a so-called, open evaluation system through the active involvement of

international experts and volunteer reviewers and also emphasised the evaluation of the quality of core research outputs through the simplification and quantification of evaluation indicators (MOST and KISTEP, 2000a : 1). For this reason, I adopted the evaluation system of the CRI programme as a case study in relation to the evaluation of R&D projects.

Meanwhile, in contrast to the UK evaluations that are carried out every two years, every four years and every five years, depending on their evaluation purposes and characteristics, all Korean evaluation exercises, in principle, are conducted once a year. So, I have chosen recently completed evaluations as examples for the case studies. These are summarised in <Table 5-5>.

**<Table 5-5> Recent Evaluation Exercises in S&T and R&D Sectors (Korea)**

Type of Evaluation		Recent Evaluation Exercises
Evaluation of S&T and R&D Policy		- Evaluation of Major Policy Issues of the MOST by the OPM/the OGPC (2001)
Evaluation of R&D Programme		- Evaluation of National R&D Programmes by the NSTC (2001)
Evaluation of Research Institution	Evaluation of Research Council	- Evaluation of Five Research Councils by the OGPC (2001)
	Evaluation of GRI	- Evaluation of Member Research Institutes by the KOCI (2002)
Evaluation of R&D Project		- Evaluation of the Creative Research Initiatives (CRI) by the MOST (2000)

In this section, I shall analyse and describe the five evaluation exercises on the basis of the four main components and 11 sub-components of the meta-evaluation model suggested in Chapter 3. In each case, the statutes concerned, a manual or an operating plan for evaluation, an evaluation report, related documents from the websites and the results of interviews with related staff have been used as basic materials for analysis.

### ***5.3.1. Evaluation of S&T and R&D Policy : Evaluation of Major Policy Issues of the MOST by the OPM/the OGPC (2001)***

In early 2002, the OPM/the OGPC conducted an evaluation of the performance of central and local government in 2001. This evaluation included major policy issues, capabilities

in implementing them and the survey of citizens' satisfaction with them of 40 central administrative offices and 16 large-unit local governments. For this evaluation, both the OGPC and the PEC issued *the Guidance for Government Performance Evaluation in 2001* in advance in January 2001. In particular, the evaluation of major policy issues of central administrative offices was conducted by the "Policy Evaluation Committee" (PEC) all of whose members are private sector experts. As a part of this work, the OGPC evaluated some policy issues that are relevant to S&T and R&D affairs in the related ministries including the MOST, the MOCIE and the MIC. Of these, the evaluation of major policy issues of the MOST directly concerns S&T and R&D policies at the national level. In what follows, after an outline of the evaluation practices of government performance in 2001, I shall deal with the evaluation exercise conducted in relation to the S&T and R&D policies of the MOST.

#### ***5.3.1.1. Overview of Government Performance Evaluation in 2001***

As a part of the evaluation of central administrative offices, the OGPC and the PEC evaluated 63 major policy issues in central government. To carry out this evaluation, the OGPC selected in advance two or three policies, which were considered the most important and representative tasks, from organisational missions of each individual ministry (OGPC and PEC, 2001b : 2-3). The major policy issues of central government fall into four groups as follows : (1) economic affairs, (2) the affairs of unification, foreign and national security, (3) general administrative affairs and (4) social and cultural affairs. Of these, the evaluation of economic affairs includes 34 major policy issues of 22 central administrative offices including the MOST (OGPC and PEC, 2002a : 8).

In the 2001 exercise, major policy issues in relation to S&T and R&D affairs include some policies of the MOST, the MOCIE and the MIC which are main actors in the S&T and R&D sectors. These policies include (1) the strategic development of leading edge technologies and (2) the development of nuclear technology and strengthening its security by the MOST, (3) the promotion of a high-valued type industrial structure by the MOCIE, (4) the utilisation of the infrastructure based on knowledge and information and (5) the promotion of information and telecommunication industry by the MIC (OGPC and PEC, 2002a : 50-61). Therefore, the evaluation of two policies of the MOST, that is, the strategic development of leading edge technologies and the development of nuclear

technology and strengthening its security, is considered typical of the evaluations of S&T and R&D policy in Korea.

Therefore, further discussion of the evaluation of S&T and R&D policies will focus on “the evaluation of the strategic development of leading edge technologies”, one of the two major policy issues of the MOST.

### ***5.3.1.2. Evaluation Paradigm***

#### **(1) Evaluation Purpose**

In general, the evaluation of government performance is conducted in order (1) to improve the effectiveness of the policy formation and its implementation in each individual ministry, (2) to ensure the accountability of the pertinent ministry for them and (3) to enhance the credibility of the government policies as a whole in the eyes of the citizens (OPM, 2001a). Such an evaluation purpose can be equally applicable to the evaluation of major policy issues of the central government, which one of the important areas of government performance evaluation.

From the general remit of the evaluation of government performance, we can draw some main purposes of the evaluation of major policy issues and, in particular, of the evaluation of the strategic development of leading edge technologies by the MOST in 2001. In this evaluation, as in most policy evaluations, both “accountability” and “policy improvement” were recognised as main evaluation purposes (Goldenberg, 1983 ; Arvidsson, 1986 : 627 ; Lee & Sampson, 1990 : 157 ; Hudson, Mayne and Thomlison, 1992 : 5), even though this was not announced explicitly in either the evaluation guidance or the evaluation report. For the former purpose, the evaluator, namely the PEC, examined the extent to which the pertinent policy had been answerable to the public interest and the needs of the public. Meanwhile, for the latter purpose, it examined how rational the policy formulated had been and how effectively it had been implemented (OGPC and PEC, 2002b : 1, 4).

#### **(2) Evaluation Object**

This evaluation selected “the strategic development of leading edge technologies”, one of the major policy issues of the MOST, as the evaluation object. This evaluation object



comprised four sub programmes and 17 projects. The four sub programmes were as follows : (1) the establishment of infrastructure for the promotion of the biotechnology industry, (2) the full-scale development of state-of-the-art nanotechnology, (3) the development of space and aeronautics technology for peaceful utilisation, (4) the transformation of conventional industry into new knowledge-based industry and (5) the development of technology for the improvement of energy utilisation. Each sub programme consisted of between one and six projects (OGPC and PEC, 2002b : 2-3).

Even though this evaluation was initially intended to judge a certain policy, namely the strategic development of leading edge technologies, it appears that the level of the evaluation object included both “sub programmes” and their “main projects” which were components of the policy. In particular, the PEC intensively investigated the extent to which the goal of the project had been achieved by reference to the initial plan set up in advance. On the other hand, the scope of the evaluation object was confined to the MOST’s activities, although other ministries were involved or interested in S&T and R&D affairs. In consequence, it seems that the evaluation object dealt with specific sub programmes and more detailed projects that were consisted of related policy of the MOST.

### **(3) Evaluation Type**

The government performance evaluation is basically conducted twice a year. In the first half of the year, the rationality of policy formulation and the effectiveness of its implementation are examined, whereas in the second half, the achievement of the policy objectives and their effect on the public are vigorously investigated (OGPC and PEC, 2001b : 6).

From this, we know that this evaluation explicitly intended to carry out an “interim” evaluation, an “output” evaluation and, where appropriate, an “impact” evaluation. In order to identify some problems in the stage of policy formulation and implementation, it needed to carry out an interim evaluation, whereas, to judge the rate and extent of the achievement of the policy objectives, an output evaluation was conducted. In addition, an impact evaluation was performed to investigate both the direct and indirect effect of the pertinent policy on the public and society as a whole.

### **5.3.1.3. Evaluation Resources**

#### **(1) Evaluation Manpower**

In general, two groups of manpower are involved in the evaluation of government performance. Firstly, the staff of “the Assistant Minister’s Office (AMO) for Policy Analysis and Evaluation”, one of the five assistant ministers in the OGPC, prepares the master plan, supports the PEC in conducting actual evaluation and coordinates evaluation results. As at April 2002, the total number of evaluation-related staff in this office was 35, and their missions were to handle diverse evaluation functions at the working level. Secondly, the members of “the PEC”, with the assistance of the experts from the support team, conducted the actual evaluation in each area under their responsibility. The PEC was composed of 30 members, including 29 non-government experts representing academia, industry, the media and other private organisations and the Minister of the OGPC. The “support team”, whose mission is to assist the evaluation activities of the PEC with high expertise in each area, comprised 29 experts mainly from academia and related research institutes (PEC, 2002b).

A total of four people were involved in the evaluation exercise in relation to the major policy issues of the MOST : one member of the PEC, one expert from the support team and two officials of the Assistant Minister’s Office (AMO) for Policy Analysis and Evaluation (OGPC and PEC, 2002b).

As a result, it is apparent that the evaluation of the MOST’s policy issues in 2001 was conducted through collaborative work between officials of the OGPC, in particular, the Assistant Minister’s Office for Policy Analysis and Evaluation, and members of the PEC and its support team. However, while internal manpower from the OGPC carried out this evaluation as their permanent work, the external manpower did this as temporary work for a restricted period.

#### **(2) Evaluation Unit**

The AMO for Policy Analysis and Evaluation under the OGPC had the principal responsibility for this evaluation. This office is in charge of many evaluation functions, such as the establishment of annual basic plans and guidelines for government performance evaluation, the development of criteria and indicators for evaluations, the

designation of duties to be evaluated and the verification of evaluations by the visiting of actual sites. Of the three director-generals in this office, two are directly engaged in the evaluation. As of April 2002, the "Director-General I for Policy Analysis and Evaluation" comprised 21 members of staff and was responsible for the evaluation of the economic policies implemented by 22 central administrative offices, whereas the "Director-General II for Policy Analysis and Evaluation" consisted of 14 members of staff and was answerable for the evaluation of non economic policies performed by another 21 central administrative offices (OPM, 2002a ; PEC, 2002a). Therefore, the evaluation of the strategic development of leading edge technologies, a major policy issue of the MOST, was under the responsibility of the Director-General I.

Consequently, the Assistant Minister's Office for Policy Analysis and Evaluation and, in particular, Director-General I for Policy Analysis and Evaluation (PAE) played a role as the evaluation unit in the evaluation of major policy issues of the MOST in 2001, a typical type of policy evaluation in the S&T and R&D sector at national level.

### **(3) Evaluation Information**

In the evaluation of major policy issues, important information was drawn from three kinds of sources. One was the formal document submitted in advance by each ministry including the operating plan of major policy issues and the self-evaluation report. By early March every year, central administrative offices are required to submit the operating plan for the major policy issues that are supposed to be evaluated. This is expected to contain (1) the outline of the policy, (2) the background to policy formulation and implementation (3) policy objectives and predictable effects and (4) the operating plan for sub programmes. At each sub programme level, same elements as those contained in the policy level are also required to be included (OGPC and PEC, 2001b : 27-28).

In addition, central administrative offices must conduct the self-evaluation of major policy issues adopting the same evaluation methods and criteria as those used in the PEC's exercise, for which operating plans were produced and already submitted to the OGPC. The findings of the self-evaluation are required to be submitted to the OGPC in advance of the evaluation work of the PEC. This report is expected to cover the main outputs of the policy issues so that they can be compared with the original plan set up at the beginning of the year (OGPC and PEC, 2001b : 33).

The other source was the extensive consultation that was conducted with related experts and members of staff involved in the pertinent policies (OGPC and PEC, 2001b : 4-5 ; OGPC and PEC, 2002a : 4).

Hence, in the 2001 exercise, much more use was made of qualitative, sometimes slightly subjective, information than of quantitative information in conducting the evaluation of the major policy issues of individual ministries.

#### 5.3.1.4. Evaluation Performance

##### (1) Evaluator

As already noted, this evaluation was actually conducted by the PEC, with administrative assistance from the Assistant Minister's Office. Under the PEC, there are five sub committees that cover policy areas grouped into similar types.

<Table 5-6> Members and their Organisations of the PEC and the Support Team

Organisations PEC & Support Team		A	RI	I	M	O	Total
<b>Chairperson</b>		1	-	-	-	-	1
<b>PEC</b>	Evaluation System & Management	1	-	1	-	-	2
	Economic Policies I	3	-	2	1	1	7
	Economic Policies II	4	2	1	-	-	7
	Social & Cultural Policies	4	-	1	1	-	6
	General Administrative Policies	6	-	-	-	-	6
	Minister of the OGPC	-	-	-	-	1	1
	<i>Sub-total</i>	18(62%)	2(7%)	5(17%)	2(7%)	2(7%)	29
<b>Support Team</b>	Economic Policies I	2	5	-	-	-	7
	Economic Policies II	-	7	-	-	-	7
	Social & Cultural Policies	2	6	-	-	-	8
	General Administrative Policies	2	5	-	-	-	7
	<i>Sub-total</i>	6(21%)	23(79%)	-	-	-	29
<b>Total</b>		25	25	5	2	2	58

A : Academia      RI : Research Institutes      I : Industry  
M : the Media      O : Others including Law Firms and Government

Each member of the sub committee has a responsibility for conducting the actual evaluation of major policy issues of one or two central ministries falling within his/her remit. He/she also receives the professional assistance from the “support team” which is equivalent to a sub committee (PEC, 2002b). The members of sub committees and support team and their original organisations as of April 2002 are shown in <Table 5-6>.

From this, we know that, in the evaluation of government performance in 2001, most of members of the PEC, namely actual evaluators, came from academia, whereas many members of the support team belonged to related research institutes. As a result, in general, the evaluators from academia and the research institutes played more important roles than others from industry and the media did.

Similar to the general exercise of this evaluation, one member, who belonged to the sub Committee for Economic Policies II, led the evaluation of major policy issues of the MOST, with the assistance of one expert from the support team and two members of staff of the Director-General I for PAE (OGPC and PEC, 2002b).

## **(2) Evaluation Method**

The members of the PEC firstly reviewed the operating plans and self-evaluation report submitted by individual ministries in advance, which contained outputs and results of pertinent policies compared with the plans set up at the beginning of the year. Next, in order to confirm the contents of these materials, they visited each ministry or site and also conducted an intensive consultation with both the experts concerned and related officials from the pertinent ministries. In particular, the findings of the self-evaluation by each ministry were directly reflected in the evaluation process by the member of the PEC (OGPC and PEC 2001b : 5).

Therefore, in the evaluation of major policy issues in 2001, a qualitative method, including in-depth consultation and a literature review in the form of the review of self-evaluation report of individual ministries, was adopted as the main evaluation method.

Meanwhile, regarding the grading methods of sub programmes and projects, each project of the sub programme is firstly given a grade on a three-tier grading system of good, fair

and unsatisfactory in a qualitative and descriptive way. Then, on the basis of the evaluation results of detailed projects, individual sub programmes are also given a certain grade on a four-tier grading system of very good, good, fair and unsatisfactory in a qualitative way. Finally, these results are integrated at the policy level to decide how well the pertinent policy has been implemented and progressed by comparison with the initial plan (OGPC and PEC, 2001a : 25-30).

### (3) Evaluation Criteria and Items

In this evaluation, in order to judge the rationality and appropriateness of the policy formulation, policy implementation and policy achievement, the evaluator, the member of the PEC, adopted the common evaluation criteria and items shown in <Table 5-7>.

<Table 5-7> Evaluation Criteria and Items for Major Policy Issues in 2001 Exercise

Stage of Policy Process	Evaluation Criteria	Evaluation Items
<b>Policy Formulation</b>	1. Validity of Objectives	1. Are the overall objectives and the objectives for the year clearly stated? 2. Do the objectives suit the higher-level government policy guidelines and do they meet situational changes?
	2. Integrity of Contents	3. Are there sufficient lower-level policy projects in support of the policy objectives? 4. Has public opinion been ascertained and have related procedures been sufficiently implemented to establish the plan?
<b>Policy Implementation</b>	3. Efficiency in Implementation	5. Are programmes being promoted in line with particular plans? 6. Are resources being used efficiently to achieve the desired results?
	4. Appropriateness of Implementation Process	7. Have changes in administrative conditions and situations been adequately grasped and taken into account in the process of implementation? 8. Are the public and interested groups being kept properly informed?
<b>Policy Achievement</b>	5. Rate of Achievement of Objectives	9. Have the initial objectives of the policy been achieved?
	6. Policy Effectiveness	10. From the public's point of view, has the effectiveness of the policy manifested itself?

These six criteria and 10 items were applied to the evaluation of all policy issues in central administrative offices. Meanwhile, in principle, individual ministries were asked to adopt these common evaluation criteria for their self-evaluation, even though they were allowed to choose some extra evaluation criteria, if they found it necessary to do so (OGPC and PEC, 2001b : 6 ; PEC, 2002a).

Thus, in this evaluation, evaluation criteria and items adopted were mainly qualitative in nature. However, there was no clear identification of the evaluation indicators in the related documents of the OGPC and the PEC.

### **5.3.1.5. *Evaluation Utilisation***

#### **(1) Evaluation Report**

The findings of government performance evaluation are reported in the presentation conference for government evaluation on a six monthly basis. In addition, they are also open to the public and submitted to the National Assembly. In the case of the 2001 exercise, in July 2001, the Prime Minister presided over a conference attended by heads of central administrative offices and members of the PEC. The conference focused on the results of the evaluations of each ministry's major policy issues and policy promotion capabilities. In January 2002, the President presided over a conference, attended by heads of central administrative offices, members of the PEC and governors of large-unit local governments. It related to a comprehensive report of the evaluation of major policy issues, institutional capabilities in implementing them, the survey of citizen's satisfaction with them in central government as well as local governments' performance (OGPC and PEC, 2001a : 23 ; PEC, 2002a).

The final report of the government performance evaluation was also published in January 2002 and can be found on the website of both the OPM (<http://www.opm.go.kr/home/40/402.html>) and the PEC (<http://www.pec.go.kr/destiny/servlet/opc/login>). This final report consisted of a comprehensive report and many individual reports for each ministry. While the former contained a general background, the summary of each ministry's evaluations, the feedback of the evaluation findings and so on, the latter covered more detailed evaluation components including evaluators, outlines of policy issues including policy objectives and main contents, evaluation findings based on evaluation criteria and the

utilisation of evaluation findings (OGPC and PEC, 2002a).

The evaluation report of major policy issues of the MOST has a similar structure and the contents to that in other cases. In particular, for the evaluation of the strategic development of leading edge technologies, this report included individual evaluation findings of five sub programmes in a descriptive way, according to six evaluation criteria. As already mentioned, these were made on the basis of the evaluation results of the detailed projects, which have been evaluated into three categories : good, fair and unsatisfactory (OGPC and PEC, 2002b : 9-24).

From this, it appears that the final report of the government performance evaluation in 2001, consisting of both the comprehensive report and individual ones, focused more on the results of the evaluation exercise. For this reason, it seems that they lacked some important components of the evaluation, including an evaluation purpose, evaluation types, evaluation information utilised, evaluation methods, an evaluation process and so on, which are normally expected to be included in an evaluation report.

## **(2) Utilisation Type**

In general, the evaluation findings of government performance are utilised for the following basic purposes : (1) giving awards to excellent institutions and officials, (2) improvement of the policy through the implementation of corrective measures and the conduct of audit and (3) the adjustment of the budget allocation for the following year (OPM, 2001a ; PEC, 2002a).

Ministries and officials found by government evaluations to have excellent performance are provided with incentives, such as awards and preferential treatment for personnel or budgeting. In the 2001 exercise, four central administrative offices were given awards (OGPC and PEC, 2002a : 45).

On the basis of the evaluation findings, the OGPC specifies corrective measures and instructs the heads of the related central administrative offices to implement them. When the heads have received instructions to implement corrections, they must immediately set up an implementation plan of corrective measures and reflect it in operating plans and strategies for the following year. They also report biannually on the status of the



implementation of the plan of corrective measures. The OGPC verifies and oversees each ministry's implementation. If necessary, the sites are visited without notice for verification. The results of verification are reflected in the self-evaluation conducted in the following year. In the 2001 evaluation exercise, a total of 323 projects in the area of major policy issues were identified as needing corrective measures. In the evaluation of the MOST's policy issue, namely the strategic development of leading edge technologies, a total of five among 17 detailed policy projects were identified for corrective measures. These were required to be reflected in operating plans for the following year (OGPC and PEC, 2002a : 44 ; OGPC and PEC, 2002c : 46-47).

In addition, the OGPC notifies the MPB and the ministries about the projects, which need to be subject to budgetary measures based on the evaluation findings. In principle, there is an effort to reflect these as far as possible in the budget process (PEC, 2002a).

Therefore, it can be said that the findings of the evaluation of government performance were normally utilised for short-term policy improvements, rather than long-term policy development. In particular, regarding the evaluation of major policy issues, the evaluation findings were considered as basic and important information used to identify some problems at the stage of both policy formulation and implementation as well as to suggest corrective measures in the future to improve sub programmes and detailed projects of the pertinent policy.

### ***5.3.2. Evaluation of R&D Programme : Evaluation of National R&D Programmes by the NSTC (2001)***

Between February and June 2001, as a part of that year's analysis, evaluation and pre-budget review of national R&D programmes, the NSTC conducted the evaluation of 161 national R&D programmes that had been implemented by 18 central ministries and administrative offices in the year 2000. These programmes totalled 2,580 billion KRW. For the purposes of this evaluation, in December 2000, the MOST, the secretariat of the NSTC, published *the Master Plan for Analysis, Evaluation and Pre-Budget Review of National R&D Programmes in 2001*. This evaluation comprised two types of work that were closely connected, although their aims were slightly different : one was the

“programme evaluation” to judge the performance of R&D programmes, while the other was the “project review” to investigate the overlapping among R&D projects, a component of the programme. The findings of the evaluation were firstly adopted as very important information to improve national R&D programmes. They were also reflected in the process of the pre-budget review of national R&D programmes for 2002.

#### **5.3.2.1. Evaluation Paradigm**

##### **(1) Evaluation Purpose**

The analysis and evaluation of national R&D programmes mainly aims to improve the effectiveness of the R&D investment funded by the government. This principal purpose comprises the following sub purposes : (1) to set up the forward policy and management system through the survey and analysis of national R&D programmes from the perspective of their socio-economic purposes, S&T areas, R&D performers and so on, (2) to establish a competitive and productive R&D system and to enhance the effectiveness of R&D programmes by evaluating the management system and outputs of diverse national R&D programmes, (3) to avoid the duplication of R&D investment by identifying linkages and overlapping among national R&D projects and (4) to promote the synergy effects of the knowledge transfer among institutions responsible for R&D programmes (MOST, 2000b : 2 ; NSTC, 2001a : 2).

Therefore, in the evaluation of national R&D programmes in 2001, the Korean government intended to achieve the following three purposes : the “accountability” of R&D related organisations, the “improvement” of national R&D programmes and “knowledge sharing” for the effective R&D management in the public sector.

##### **(2) Evaluation Object**

In principle, the evaluation object included all R&D programmes supported by the national budgets and public funds including the informatisation fund of the MIC and the nuclear energy R&D fund of the MOST. However, some programmes that were not directly relevant to actual R&D activities were excluded from this exercise. These included (1) research programmes in the area of the social sciences and humanities, (2) defence-related R&D programmes classified as confidential, (3) programmes for R&D

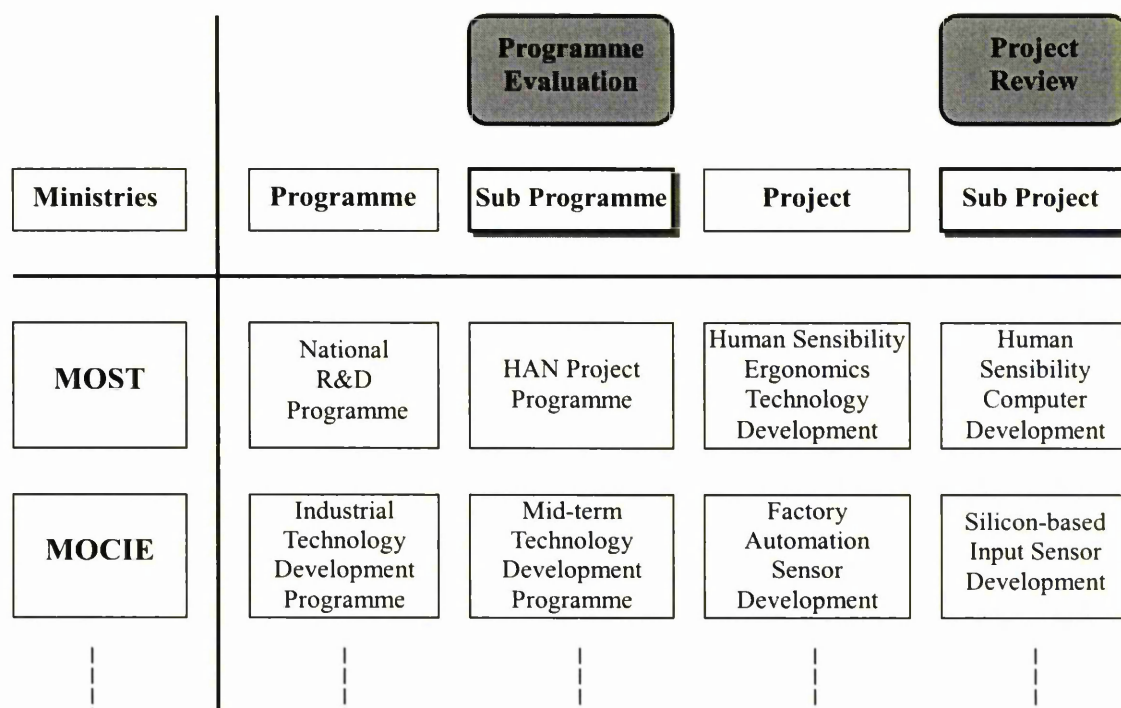
planning and management, R&D policy studies and field surveys in each ministry (4) programmes for organisational remits in R&D management institutes such as the KISTEP, the ITEP and the IITA and (5) programmes for the construction and refurbishment of research buildings. As a result, 161 R&D programmes, each representing a budget item, were selected as evaluation objects (MOST, 2000b : 6 ; KISTEP, 2002b : 94 ; Yang, 2002 : 35). Therefore, the scope of the evaluation object contained most R&D programmes at the national level.

Regarding the level of evaluation objects, actual objects were differently defined depending on the programme evaluation or project review. In the programme evaluation, the detailed evaluation objects were 161 sub-programmes and these were categorised into four main groups and 12 sub groups as follows : (1) basic research programmes, (2) publicly used technology programmes, (3) welfare technology programmes, (4) short-term industrial research programmes, (5) mid and long-term industrial research programmes, (6) human resource development programmes, (7) infrastructure building programmes, (8) international cooperative R&D programmes, (9) national research institutes (NRIs) support programmes, (10) the KRCF's basic technology support programmes, (11) the KORP's public technology support programmes and (12) the KOCI's industrial technology support programmes (MOST, 2000b : 8 ; NSTC, 2001a : 12).

On the other hand, in the project review, the evaluation objects were the 9,533 sub-projects that were the lowest units of the pertinent programmes. The R&D projects were classified into 14 S&T areas. These were (1) basic science, (2) mechanical engineering, (3) materials, (4) biotechnology, (5) agriculture, fishing and forestry, (6) health and medicine science, (7) applied chemistry and chemical engineering, (8) energy and resources, (9) nuclear energy, (10) information, electronics and communication, (11) the environment, (12) construction, (13) transportation including space technology and (14) ocean science and atmospheric science (MOST, 2000b : 8 ; NSTC, 2001a : 13). The evaluation objects at the two different evaluation levels are shown in (Figure 5-3).

As a result, it is apparent that, while the scope of the evaluation objects, in principle, included all R&D programmes performed by the central government, the object actually evaluated were at sub programme or sub project level, depending on the evaluation purposes in each case. Hence, we know that the 2001 evaluation exercise was actually

conducted at the level of sub programmes or sub projects, even though its purpose was to examine the overall performance of programmes themselves.



Source : Yang (2002 : 31)

(Figure 5-3) Objects of the Programme Evaluation and the Project Review

### (3) Evaluation Type

As already discussed, this evaluation was undertaken to improve the effectiveness of R&D investment made by the government. In order to achieve the principal evaluation purpose, two different evaluations were conducted. One was to judge the performance of diverse R&D programmes. The other was to review the overlap among R&D projects and to identify linkages among them (NSTC, 2002a).

Therefore, it is apparent that the NSTC's evaluation took a short term perspective and focused a great deal on both "interim" and "output" evaluation. Thus far, no impact evaluation has been conducted by the NSTC, even though it is considered one of the main evaluation purposes of this evaluation.

### **5.3.2.2. Evaluation Resources**

#### **(1) Evaluation Manpower**

Three groups of manpower were involved in the NSTC's evaluation. The first was the members of the evaluation committee, which was entirely made up of external experts, and the second was the internal members of staff of the MOST and the KISTEP. In addition, there was a supporting team to assist the committees during the evaluation exercise.

There were three types of evaluation committee : the steering committee, the evaluation committees for sub R&D programmes and the review committees for R&D projects. These committees conducted the actual evaluation of R&D programmes and R&D projects and coordinated the findings of each piece of evaluation work. "The Steering Committee" comprised 19 external members including 4 private members of the operating committee of the NSTC, two chairpersons of sub committees of the NSTC and 11 chairpersons of evaluation committees for sub programmes.<sup>10</sup> Its mission was to review and coordinate the evaluation results from 14 review committees for R&D projects and eight evaluation committees for sub programmes. "The Evaluation Committee for Sub Programmes" comprised eight sub committees reflecting the characteristics of R&D programmes, although these were actually categorised into 12 groups. Individual sub committees covered such programmes as : (1) basic research programmes, (2) publicly used technology programmes, (3) welfare technology programmes, (4) short-term industrial research programmes, (5) mid and long-term industrial research programmes, (6) human resource development and international cooperative R&D programmes, (7) infrastructure building programmes and (8) national research institutes (NRIs) and GRIs support programmes. Each committee consisted of around 10 experts and a total of 83 experts were involved in these committees. "The Review Committee for R&D Projects" was also divided into 14 sub committees according to the S&T areas already noted above. Each committee comprised 10 experts and a total of 140 experts were involved in these committees (NSTC, 2002a : 325-350). As a result, a total of 242 members sitting on the three types of committee were engaged

---

<sup>10</sup> In the actual evaluation exercise, some groups of R&D programme were integrated into one evaluation committee, even though they were classified into 12 groups. For this reason, fewer sub evaluation committees than programme groups were organised in the 2001 exercise.

in the evaluation of 161 sub R&D programmes and the review of 9,533 R&D projects.

Some members of staff in the MOST and the KISTEP were also involved in the evaluation, as the secretariat of the NSTC and of the evaluation committee. In the MOST, “the Science and Technology Assessment Division” was responsible for this evaluation through the preparation of a master plan for the analysis, evaluation and pre-budget review of national R&D programmes every year. As of April 2002, this division comprised eight officials (MOST, 2002b). On the other hand, in the KISTEP, “the Division of National Research and Development Evaluation” is in charge of the overall evaluation work at the operating level. As of April 2002, the division consisted of 21 members of staff (KISTEP, 2002b). Thus, in the 2001 exercise, approximately 30 internal members of staff from the MOST and the KISTEP were involved.

Meanwhile, the supporting team was organised in order to provide many committees with administrative assistance. This team comprised 16 members, whose were R&D management experts from central administrative offices and their affiliated institutions (NSTC, 2002a : 351).

In sum, around 270 members, including 240 external experts and 30 internal staff, were fully or partially engaged in the evaluation of national R&D programmes in 2001 for a period of approximately four months.

## **(2) Evaluation Unit**

Both “the Science and Technology Assessment Division” of the MOST and “the Division of National Research and Development Evaluation” of the KISTEP acted as evaluation units. While the role of the former was to set up the master plan for the analysis, evaluation and pre-budget review of national R&D programmes, the latter’s role was to implement this plan and to assist the evaluation activities of the committees (MOST, 2002b ; KISTEP, 2002b). In particular, based on the Framework Act on Science and Technology, the KISTEP was designated as the secretariat of the evaluation committee for national R&D programmes (MOST, 2001a). As already noted, there were 21 members of staff in the Division of National Research and Development Evaluation of the KISTEP. Of these, four members of staff devoted themselves fully to this evaluation. As a result, two divisions acted together as the evaluation unit, even though they were also

responsible for other matters in each organisation.

### **(3) Evaluation Information**

In this evaluation, most of the important information came from two kinds of source : (1) formal documents submitted by individual ministries and (2) presentations made by the programme managers. Each central administrative office had to submit the required materials to the MOST that was the NSTC secretariat by the end of February. The materials were expected to contain the overview of the programme, its current status, the main performance and outcomes and detailed project lists including their research results. In addition, each programme manager was required to give a presentation of the research outputs in order to supplement the related materials (NSTC, 2001a : 2-3 ; Yang, 2002 : 18).

Consequently, diverse qualitative information was utilised as the main source for this evaluation. However, it seems that some rather infrequent use may have been made of quantitative information.

#### ***5.3.2.3. Evaluation Performance***

##### **(1) Evaluator**

As already described, three groups of evaluation manpower were involved in this evaluation. Of these, it was the members of two types of evaluation committee, namely the evaluation committee for sub R&D programmes and the review committee for R&D projects who conducted the actual evaluation work. All the members of the two committees came from academia, the research institutes and industry and were recommended in advance by individual ministries to the NSTC secretariat.

“The Evaluation Committee for sub R&D Programmes” was divided into eight sub committees, each of which comprised 10 to 13 experts. “The Review Committee for R&D Projects” consisted of 14 sub committees, one for each S&T area concerned, and each sub committee comprised 10 external professionals. The detailed breakdown of the membership of related sub committees is shown in <Table 5-8> and <Table 5-9>.

Therefore, we know that there were more members from academia than from industry

and research institutes, although their proportion was not much higher. So, we may see that this evaluation was entirely conducted by external evaluators, even though there was much assistance from members of staff in the KISTEP and the supporting team.

**<Table 5-8> Membership of the Evaluation Committee for sub R&D Programmes**

<b>Institution Sub Committee</b>	<b>Industry</b>	<b>Academia</b>	<b>Research Institutes</b>	<b>Total</b>
Basic Research	1	4	5	10
Publicly-used Technology	1	5	4	10
Welfare Technology	2	5	3	10
Short-term Industrial Technology	4	2	4	10
Mid and Long-term Industrial Technology	1	4	5	10
Human Resource Development / International cooperative R&D	2	5	3	10
Infrastructure Building	3	4	3	10
NRIs & GRIs Support	6	6	1	13
<b>Total</b>	<b>20 (24%)</b>	<b>35 (42%)</b>	<b>28 (34%)</b>	<b>83 (100%)</b>

**<Table 5-9> Membership of the Review Committee for R&D Projects**

<b>Institution Sub Committee</b>	<b>Industry</b>	<b>Academia</b>	<b>Research Institutes</b>	<b>Total</b>
Basic Science	1	5	4	10
Mechanical Engineering	2	5	3	10
Materials	4	3	3	10
Biotechnology	3	5	2	10
Agriculture, Fishing & Forestry	3	4	3	10
Health and Medicine Science	1	7	2	10
Applied Chemistry & Chemical Engineering	2	5	3	10
Energy and Resource	2	3	5	10
Nuclear Energy	4	4	2	10
Information, Electronic and Communication	4	3	3	10
Environment	4	3	3	10
Construction	2	6	2	10
Transportation (including Space Technology)	2	5	3	10
Ocean Science and Atmospheric Science	3	4	3	10
<b>Total</b>	<b>37 (27%)</b>	<b>62 (44%)</b>	<b>41 (29%)</b>	<b>223 (100%)</b>



## **(2) Evaluation Method**

In this evaluation, a literature review was adopted as the principal evaluation method. Thus, in order to judge the performance of R&D programmes and the effectiveness of their management systems, the evaluators, namely the members of the pertinent committees, reviewed official documents submitted by individual central ministries and agencies in advance of the evaluation exercise (NSTC, 2001a : 14-15). Therefore, it appears that, in the evaluation of national R&D programmes in 2001, a qualitative analytical approach, in the form of a literature review, was the principal method adopted.

The evaluators have to give a certain grade to each programme on a five-tier grading system that comprised excellent (A grade), good (B grade), fair (C grade), unsatisfactory (D grade) and need for improvement (E grade). Normally, approximately 10%, 20%, 40%, 20% and 10% of programmes are graded A, B, C, D and E, respectively. In principle, there was no exception for this grading system. As a result, in the 2001 exercise, of all 161 R&D programmes, there were 17 programmes of A grade (10.6%), 31 programmes of B grade (19.2%), 65 programmes of C grade (40.4%), 31 programmes of D grade (19.2 %) and 17 programmes of E grade (10.6%) (NSTC, 2001a : 15, 94 ; Yang, 2002 : 42).

## **(3) Evaluation Criteria and Indicator**

The evaluation criteria for programme evaluation were slightly different from those for project review. In the programme evaluation, three main criteria were adopted : “the appropriateness of the programme objectives”, “the efficiency of implementation” and “the level of performance and outputs”. Each of these main criteria was again divided into three to six sub criteria. Some extra criteria were also used in the programme evaluation. These included “the need for adjusting the programme’s scope”, “integration and/or linkage with other programmes” and “re-planning of the programme”. The weights for the evaluation criteria were differed according to the nature and characteristics of programme groups. Each criterion was given a certain grade on a five-tier grading system and it was also expected that the reason for that grade and any recommendations for future improvement would be presented in a qualitative way (NSTC, 2001a : 14-15 ; Yang, 2002 : 35).

In the project review, “the redundancy of R&D projects” and “the need for integration and/or linkage with other projects” were adopted as the major evaluation criteria (NSTC,

2001a : 16).

Thus, it appears that qualitative evaluation criteria were adopted to a large extent in the 2001 exercise. Meanwhile, there was no clear identification of the evaluation indicators in the related official documents such as the MOST's master plan for this evaluation and the evaluation report produced by the NSTC.

#### **5.3.2.4. Evaluation Utilisation**

##### **(1) Evaluation Report**

The evaluation report for this exercise was published in July 2001 and it is also available on the website of both the NSTC (<http://www.nstc.go.kr>) and the MOST (<http://www.most.go.kr>). The report contained some important components including the evaluation purposes, the evaluation objects, evaluators, the evaluation methods, the evaluation findings and recommendations. In particular, many parts of the report were devoted to the discussion of evaluation findings and major recommendations in individual sub R&D programmes and R&D projects (NSTC, 2001a).

##### **(2) Utilisation Type**

Through this evaluation, the Korean government intends to enhance the effectiveness of the R&D investment as well as to improve the planning and management systems of all national R&D programmes. Firstly, the evaluation findings of the R&D programmes were directly reflected in the allocation of research funds to national R&D programmes for 2002 through the process of pre-budget review. These findings were notified to the Prime Minister and the Minister of the MPB (MOST, 2001a ; NSTC, 2001 : 2).

On the basis of the evaluation findings, some major recommendations for the improvement of national R&D programmes at national level were suggested. These were divided into seven detailed groups as follows : (1) re-planning, (2) re-configuring implementation strategy, (3) linkages with other programmes, (4) expanding support, (5) focusing on high priority research fields, (6) integration with other programmes and (7) others. Consequently, a total of 70 programmes were identified as needing improvement in 2002 (NSTC, 2001 : 99-100, 110 ; Yang, 2002 : 42, 45). The detailed recommendations are shown in <Table 5-10>. From these, it appears that, among the eight programme

groups, considerable improvements were made in the group of welfare technology development programmes and were made to 17 sub programmes. Meanwhile, of the seven types of recommendation, “re-planning” of the pertinent sub programme was chosen much more than other types and was given in relation to 19 programmes.

<Table 5-10> Recommendations for Improvement of R&D Programmes

Recom. Groups	Re- Planning	Re- Configuring Imple- mentation Strategy	Linkage with other Programmes	Expanding Programme	Focusing on High Priority Research Fields	Integration with other Programmes	Others	Total
Basic Research	-	-	-	-	-	-	-	-
Publicly Used Technology	3	-	2	-	-	-	3	8 (12%)
Welfare Technology	4	-	5	2	4	-	2	17 (24%)
Short Term Industrial R.	-	1	2	2	1	-	1	7 (10%)
Mid & Long Term Industrial R.	1	-	2	3	-	-	2	8 (12%)
International Cooperation	-	1	-	1	-	-	-	2 (3%)
Human Resource D.	1	-	-	-	-	-	-	1 (1%)
Infra. Building	4	-	1	2	1	-	2	10 (14%)
NRIs Support	2	1	2	1	1	-	3	10 (14%)
GRIs Support	4	-	1	-	-	1	1	7 (10%)
<b>Total</b>	<b>19 (27%)</b>	<b>3 (4%)</b>	<b>15 (22%)</b>	<b>11 (16%)</b>	<b>7 (10%)</b>	<b>1 (1%)</b>	<b>14 (20%)</b>	<b>70 (100%)</b>

Source : Yang (2002 : 45)

On the other hand, two types of recommendation were made, concerning the improvement of R&D projects. While a total of 43 R&D projects were seen as redundant investment, a total of 121 R&D projects needed to be linked with other projects (NSTC, 2001a : 110 ; Yang, 2002 : 47). The concrete recommendations in relation to the S&T areas are shown in <Table 5-11>.

Therefore, it can be said that the evaluation findings of national R&D programmes were mainly utilised for the short-term improvement for sub R&D programmes and projects, rather than the long-term development of R&D programmes themselves and the related

S&T or R&D policies. In particular, among the diverse utilisations of the evaluation, more emphasis was placed on the allocation of research funds among national R&D programmes for the following year, which is recognised as one of the most powerful and effective policy means.

**<Table 5-11> Recommendations for Improvement of R&D Projects**

<b>S&amp;T Area</b>	<b>Total R&amp;D Projects</b>	<b>Redundant Investment among Projects</b>	<b>Need for Linkage with other Projects</b>
Basic Science	980	2	10
Mechanical Engineering	983	3	6
Materials	742	-	10
Biotechnology	518	-	2
Agriculture/Fishing/Forestry	1,181	-	2
Health & Medical Science	769	4	32
Applied Chemistry & Chemical Engineering	613	2	17
Energy & Resource	394	-	6
Nuclear Energy	193	-	4
Environment	440	6	2
Construction	170	-	-
Electronics and IT	2,083	16	17
Transport	228	-	2
Atmosphere & Ocean Science	112	2	6
Others	-	8	5
<b>Total</b>	<b>9,533</b>	<b>43 (0.5%)</b>	<b>121 (1.3%)</b>

Source : NSTC (2001a : 110)

### ***5.3.3. Evaluation of Research Council : Evaluation of Five Research Councils by the OGPC (2001)***

Between May and June 2001, the OGPC evaluated the performances and plans of five research councils implemented in 2000. This exercise was the second since the Korean research council system had been introduced in 1999. The OGPC circulated the evaluation manual for this work to the research councils concerned in April 2001. As already mentioned, the OGPC is responsible for the supervision of research councils and their member research institutes, which are the principal actors in public R&D activities including those in the social sciences. The findings of this evaluation were notified to the

Minister of the MPB and those findings in relation to the three S&T related research councils were also notified to the NSTC.

#### **5.3.3.1. Evaluation Paradigm**

##### **(1) Evaluation Purpose**

The evaluation purpose of this exercise can be clearly drawn from official documents such as the evaluation manual and the evaluation report. This evaluation was conducted with the intention of both maximising the effects of the research councils' activities and ensuring the legitimacy and the effectiveness of their management, in order to develop their member research institutes so that they possessed a world-class competitiveness (OGPC, 2001a : 2 ; ECRC and KISTEP, 2001 : 4).

From these official statements, it is apparent that the OGPC conducted this evaluation, in order to secure both the "accountability" and the "improvement" of the management of research councils.

##### **(2) Evaluation Object**

Five research councils were included in this exercise, even though their supporting research areas were slightly different. These research councils were the Korea Council of Economic and Social Research Institutes (KCESSRI), the Korea Council of Humanities and Social Research Institutes (KCHS), the Korea Research Council of Fundamental Science and Technology (KRCF), the Korea Research Council for Industrial Science and Technology (KOI) and the Korea Research Council of Public Science and Technology (KORP) (OGPC, 2001a : 3 ; ECRC and KISTEP, 2001 : 5).

In order to evaluate the research councils, three types of evaluation object were adopted : (1) the research council's own activities and the mission of supporting their member GRIs, (2) their internal management systems and (3) the satisfaction of related institutions. Of these, the third evaluation object was not directly reflected in the final evaluation results, although it was utilised as useful reference information for the evaluation of the first and the second objects. Each evaluation object was given a different weight according to its importance from the perspective of the evaluation purposes. 70% of the weight fell on the research councils' own activities and the mission of supporting their GRIs, whereas 30%

fell on the internal management system. The two evaluation objects were again divided into more specific sub objects, reflecting the evaluation purposes. The evaluation of each object only covered the performance and the plan established in relation to it during the year 2000. (OPM, 2001c ; ECRC and KISTEP, 2001 : 5-7). The detailed evaluation objects and their weight in the 2001 evaluation exercise are shown in <Table 5-12>.

**<Table 5-12> Main and Sub Evaluation Objects for Research Councils**

<b>Main Evaluation Object</b>	<b>Sub Evaluation Object</b>
<b>Research Council's Own Activities &amp; Mission of Supporting Member GRIs (70%)</b>	Appropriateness of R&D Planning and Development Plan for Member Institutes (8)
	Support for Specialisation of Research Areas in Institutes & Coordination of their Research Functions (12)
	Evaluation System for Member Institutes (18)
	Support and Management of R&D Activities (22)
	Management of Administrative Affairs (10)
<b>Internal Management System (30%)</b>	Operation of Board of Directors (6)
	Operation of Planning and Evaluation Committee (6)
	Administrative Affairs of General Secretary (13)
	Correction to Reflect the Evaluation Findings in the Previous Year (5)
<b>Others (0%)</b>	Satisfaction of Related Institutions (0)

\* *Figures in parentheses represent the weight given to each evaluation object.*

Thus, it appears that the evaluation of the five research councils focused mainly on managerial affairs at the operating level, including the management of R&D activities carried out by their member research institutes. In particular, more weight was placed on the effectiveness of the mission to support member research institutes rather than on the appropriateness of their internal management system.

### **(3) Evaluation Type**

There is not an explicit announcement in relation to the evaluation type within the official documents. However, it can be inferred from the evaluation objects mentioned above. In this evaluation, an “output” evaluation and “assessment” of the R&D planning and the development plan for member institutes were conducted simultaneously, even though the

output evaluation was given great emphasis. The judgement of past performance in relation to internal and external management systems was a typical output evaluation, whereas the prospective investigation of diverse plans in each research council was considered a kind of assessment task in the R&D management process.

### **5.3.3.2. *Evaluation Resources***

#### **(1) Evaluation Manpower**

Two groups of evaluation manpower were involved in the evaluation of the five research councils. These were on the one hand members of evaluation committee for research councils, which were entirely made up of external experts and, on the other hand, internal members of staff of the KISTEP and of the “Director-General for R&D Support”, one of the three director-generals in the “Assistant Minister’s Office for Regulatory Reform” under the OGPC (OGPC, 2001a ; ECRC and KISTEP, 2001).

The members of the evaluation committee conducted the actual evaluation of five research councils on the basis of the evaluation manual produced in advance by the OGPC. The committee comprised a total of 10 members, all of whom came from the private sector, including academia, the research institutes, industry and the media (ECRC and KISTEP, 2001 : 144).

There were two types of internal members of staff engaged in this evaluation. Firstly, the officials of the Director-General for R&D Support were involved in the evaluation as a part of the supervision of research councils and their member research institutes. Of the 13 staff of the director-general, six were directly involved in this evaluation. They prepared the master plan for this work in early 2001 and circulated the evaluation manual to the research councils in April 2001 in advance of the exercise in between May and June (OPM, 2002a). Next, some members of the “the Division of National R&D Evaluation” of the KISTEP were also involved in this exercise, on the basis of contracts commissioned by the OGPC. Of the 21 members of this division, two members of staff were directly engaged in this evaluation, acting the secretariat to the evaluation committee for the research councils (ECRC and KISTEP, 2001 : 10 ; KISTEP, 2002b).

Consequently, around 15 members were involved in the evaluation of the five research

councils in 2001, although their individual roles were different in that the external members, namely the evaluation committee, conducted the actual evaluation on the basis of the evaluation manual, whereas the internal members produced the evaluation manual or acted as the secretariat to the evaluation committee.

## **(2) Evaluation Unit**

In this evaluation, the Director-General for R&D Support under the OGPC was the evaluation unit that was responsible for organising overall evaluation. In particular, both “the Support Team of Humanities, Economic and Social Sciences Research” and “the Support Team of Science and Technology Research”, two of the four support teams under the director-general, had a principal duty of planning and supervising this evaluation exercise. As already noted, their main mission was to prepare the master plan and to issue the evaluation manual after consultation from external experts. There were three members of staff in each team (ECRC and KISTEP, 2001a : 10 ; OPM, 2002a).

## **(3) Evaluation Information**

Important information was drawn from both the materials submitted by each research councils and the questionnaire survey conducted by the committee. For this evaluation, firstly, individual research councils were required to submit to the evaluation committee a total of 33 types of materials, which were slightly different according to each sub evaluation object. While a total of 16 types of materials were requested for the evaluation of “the research council’s own activities and the mission of supporting their member GRIs”, a total of 17 types of materials were required for the evaluation of “their internal management system”. Of the 33 types of materials, five materials related to the quantitative information, whereas the rest concerned the qualitative information. The detailed information requirements were described in the evaluation manual (OGPC, 2001a : 8-33).

Secondly, the questionnaire survey also provided the committee with useful information. It covered five main categories and contained 10 questions that corresponded to some important evaluation objects. The main categories of questions related to (1) individual council’s R&D planning and development plan for their member institutes, (2) the specialisation of research fields in member institutes and coordination of their research functions, (3) the evaluation system of member research institutes, (4) the support and



management of R&D activities in member institutes and (5) the operation of the secretary general in each council.

From this, in general, more emphasis was given to qualitative than to quantitative, even though both might be required depending on the nature and characteristics of the evaluation object.

### 5.3.3.3. Evaluation Performance

#### (1) Evaluator

The actual responsibility for the evaluation of the five research councils was held by the “Evaluation Committee for Research Councils”. This committee carried out the evaluation of the research councils and prepared the evaluation report. It was chaired by a private sector professional and included 10 members from academia, the research institutes, industry and the media (ECRC and KISTEP, 2001 : 144). The detailed membership of the committee is shown in <Table 5- 13>.

<Table 5-13> Membership of the Evaluation Committee for Research Councils

Fields Committee	Academia		Research Institute	Industry	Media	Total
	Science & Engineering	Social Sciences				
<b>Chairperson</b>	-	1	-	-	-	1
<b>Members</b>	2	4	1	1	1	9
<b>Total</b>	2	5	1	1	1	10

The committee was divided into four sub committees that were in charge of detailed sub evaluation objects. These were “the Coordination Sub Committee”, “the Sub Committee of Planning and Evaluation Matters”, “the Sub Committee of Support and Management for Member Institutes” and “Sub Committee of Research Councils’ own Management System” (ECRC and KISTEP, 2001 : 10-11).

Therefore, it is apparent that, in the actual exercise, only external evaluators participated in the evaluation, even though they obtained a great deal of assistance from the internal staff of the OGPC and the KISTEP.

## **(2) Evaluation Method**

In this evaluation, the main evaluation method adopted was a qualitative one of which took the form of a literature review and a site visit. Each sub committee reviewed diverse materials submitted by individual research councils as well as the result of the questionnaire survey. They also visited all research councils for a period of one week, in order to get supplementary information (ECRC and KISTEP, 2002 : 9-11).

In addition, the committee conducted the questionnaire survey as a part of the evaluation exercise. As already described, the questionnaire comprised 10 questions, one for each main evaluation object. A total of 200 questionnaires were distributed to 30 to 60 people related to each research council. The main respondents included the president and some senior members of staff in member research institutes, two civilian members of the board of directors of each council, two members of the planning and evaluation committee of individual councils and some officials from related ministries. The result of the questionnaire was analysed in both a quantitative and qualitative way and was also utilised as useful reference information for the evaluation exercise (ECRC and KISTEP, 2001 : 125-135).

In consequent, it is apparent that, in order to conduct the evaluation of the five research councils, a literature review, a site visit and a questionnaire survey were adopted as the evaluation method. It seems that the adoption of a mainly qualitative evaluation method was directly related to both the evaluation purpose and the evaluation object, which were defined and set up in a way that was more qualitative rather than quantitative.

## **(3) Evaluation Criteria and Indicator**

Specific evaluation criteria and indicators, reflecting the attributes of each evaluation object, were adopted. Two or three evaluation criteria were adopted to judge each sub evaluation object. Then, each evaluation criterion was again divided into one to three evaluation indicators. As a result, in this evaluation excise, a total of 21 evaluation criteria and 36 evaluation indicators were adopted (ECRC and KISTEP, 2001 : 6-7). Detailed summary of the evaluation criteria and indicators adopted is shown in <Table 5-14>.

For each indicator, the evaluation was made in a qualitative way and then a certain grade was given on a five-tier grading system. However, the grading system was not used as a

relative evaluation method to investigate each council's performance in comparison with others, but as an absolute method to judge the performance of individual research councils independently without comparing others (ECRC and KISTEP, 2001 : 8).

<Table 5-14> Evaluation Criteria and Indicators for Research Councils

<b>Evaluation Object</b>	<b>Sub Evaluation Object</b>	<b>Evaluation Criteria</b>	<b>Evaluation Indicators</b>
<b>RC's own Activities &amp; Mission of Supporting Member GRIs</b>	R&D Planning and Development Plan for their Member Institutes	- completeness of contents and strategies - rationality of planning process and efforts for implementation	5
	Support of specialisation of Research Areas in Institutes & Coordination of their Research Functions	- results of specialisation of research areas in member institutes - results of coordination of research functions among institutes - rationality of process in specialisation and coordination	6
	Evaluation System for Member Institutes	- appropriateness of evaluators - rationality of system and process - utilisation of evaluation findings	6
	Support and Management of R&D Activities	- appropriateness of supporting member institutes for R&D activities - review process of budget and R&D plan - support for collaborative research	6
	Management of Administrative Affairs	- degree of support for management rationalisation in institutes - results of management rationalisation in member institutes	2
<b>Internal Management System</b>	Operation of Board of Directors	- rationality of operation of board of directors	1
	Operation of Planning & Evaluation Committee	- appropriateness of committee members - results of committee activity	2
	Administrative Affairs of General Secretary	- rationality of sub teams and staffs - soundness of budget and accounting - appropriateness of management process	6
	Correction of Findings in Previous Year	- efforts and results of correction of evaluation findings of 2000	1
<b>Others</b>	Satisfaction by Related Institutions	- satisfaction by related institutions	1
<b>Total</b>		21	36

Consequently, in this evaluation, it appears that qualitative evaluation criteria and indicators were mainly adopted, although a few quantitative ones were utilised, depending on the evaluation object.

#### **5.3.3.4. Evaluation Utilisation**

##### **(1) Evaluation Report**

In June 2001, the evaluation committee produced the evaluation report on the five research councils for 2000, with administrative assistance from the KISTEP staff. Unlike in the case of other exercises, this report is not available on the website of related institutions. This report covers evaluation components such as an evaluation purpose, evaluation objects, evaluation methods, the list of evaluators, the evaluation results of individual research councils, some recommendations, the results of the questionnaire survey and so on. In particular, in the part relating to the evaluation results of each research council, it provided an overview of each council, a synthesis of the evaluation findings, the results of the evaluation of individual indicators and the correction of unsatisfactory matters in previous year's evaluation (ECRC and KISTEP, 2001).

##### **(2) Utilisation Type**

In general, the evaluation findings are notified to the minister of the MPB. In particular, those in relation to the three research councils in S&T sector, namely the KRCF, the KORP and the KOCI, are also notified to the NSTC (OPM, 2001c ; OGPC, 2001a : 1).

However, there was no explicit announcement within the evaluation manual and the evaluation report about the detailed utilisation of evaluation findings. This could be inferred indirectly from the above mentioned evaluation purpose as well as the referential statutes including the act on establishment, operation and promotion of the government-supported research institutes (GRIs). Consequently, it is expected that the findings of the evaluation of the five research councils will be utilised as policy tools to adjust the budget for R&D activities in the public sector as well as to coordinate the national R&D programmes implemented by their member research institutes.

#### **5.3.4. Evaluation of Government-supported Research Institutes (GRIs) : *Evaluation of Member Research Institutes by the KOCI (2002)***

As already noted, the five Korean research councils have carried out the evaluation of their member research institutes, the GRIs, every year since 2000. As a part of this work,

the Korea Research Council for Industrial Science and Technology (KOCI), one of three S&T related research councils, conducted an evaluation of the performance of the seven member research institutes in 2001 and of their research plans for the next year. This evaluation was conducted during a period of around one month between early February and early March 2002. The evaluation findings were submitted to the Prime Minister and the Minister of the MPB, and are reflected in the strategies and plans of the pertinent institute's management and R&D programmes for the coming year.

#### ***5.3.4.1. Evaluation Paradigm***

##### **(1) Evaluation Purpose**

This evaluation was conducted to accomplish the following four objectives : (1) to provide member institutes with a forward direction for their developments through the analysis of their mission, function, vision and strategy, (2) to induce them to reform their internal management system actively in order to improve research productivity, (3) to enhance their capabilities continuously and actively in order to create value and (4) to disseminate their research outputs and knowledge (KOCI, 2002c : 1).

Therefore, it appears that “improving” the internal management system and research programmes of member institutes was the most important evaluation purpose. In addition, both “accountability” for the overall management in each research institute and the “dissemination of knowledge” to related public and private sectors were considered further evaluation purposes.

##### **(2) Evaluation Object**

The evaluation of the KOCI in 2002 covered seven member research institutes : the Korea Institute of Oriental Medicine (KIOM), the Korea Institute of Industrial Technology (KITT), the Electronics and Telecommunications Research Institute (ETRI), the Korea Food Research Institute (KFRI), the Korea Research Institute of Machinery and Materials (KIMM), the Korea Institute of Chemical Technology (KRICT) and the Korea Electrotechnology Research Institute (KERI) (KOCI, 2002b : 5).

The evaluation object comprised two main areas, namely R&D matters and internal managerial matters, which had been implemented between January and December 2001.

Each main evaluation object was given the same weight and was again divided into several sub evaluation objects, which were given a different weight, depending on the characteristics of individual research institutes (OPM, 2001c ; KOCI, 2002c : 9-11). The main and sub evaluation objects are summarised in <Table 5-15>.

**<Table 5-15> Main and Sub Evaluation Objects for GRIs under the KOCI**

Main Evaluation Object	Sub Evaluation Object
<b>R&amp;D Matters (50%)</b>	1. Selection System of R&D Projects (10)
	2. Management System of R&D Projects (10)
	3. Performance of R&D Project (7 to 13, depending on the institute)
	4. Specialisation of Research Field (8)
	5. Collaborative Research among Industry, Academia and other Institute (3 to 5, depending on the institute)
	6. Exploitation and Transfer of Research Outputs (4 to 10, depending on the institute)
<b>Internal Managerial Matters (50%)</b>	7. Establishment of Management Goal and its Achievement (15)
	8. Internal Organisations and Humane Resources (12)
	9. Management of Finance and Budget (10)
	10. Rationalisation of Management System (13)

\* *Figures* in parentheses represent the weight of each evaluation object.

Therefore, it seems that the level of evaluation object tended to be operational matters in the short-term perspective, rather than strategic affairs in the mid and long-term perspective. Meanwhile, in principle, the scope of the evaluation object covered the past performance of R&D activities and internal management. However, some sub evaluation objects, for example the establishment of management goal and its achievement, included the past performance as well as future plans, where appropriate.

### **(3) Evaluation Type**

As already discussed, this evaluation was mainly undertaken to ensure the effectiveness of R&D activities and internal management systems. For this reason, much emphasis was placed on the investigation of the overall past performance of each institute. In addition, a partial judgement was made of the future plans and strategies of the institute management, where appropriate (KOCI, 2002c : 9-11).

From this, it seems that the KOCI work was carried out both as a retrospective “output” evaluation and as a prospective “assessment”. Regarding the output, the KOCI considered, in equal measure, the performance of R&D activities and internal management result that had been realised.

#### **5.3.4.2. Evaluation Resources**

##### **(1) Evaluation Manpower**

Two groups of manpower were involved in the KOCI's evaluation. The first was the external experts consisting of the evaluation committee and the support team, and the second was the internal members of the KOCI.<sup>11</sup>

“The Committee for the Evaluation of Research Institutes” conducted the actual evaluation of the seven research institutes and comprised 17 external experts. Of the 17 members, 12 were the members of the planning and evaluation committee of the KOCI, while five were specialists in the areas of knowledge management, accounting, human resources and organisation theory and came from the outside. “The support team”, which was awarded the contract for this work by open competition, comprised five specialists in the area of R&D management and evaluation from the Science and Technology Policy Institute (STEPI) (KOCI, 2002c : i, 12-13). Regarding internal staff, three members of staff in “the Evaluation and Management Team”, one of two teams in the KOCI, were directly involved in this evaluation. Their role was to draw up a evaluation plan and to

---

<sup>11</sup> The Korean research councils have a great of similarity in their organisational structure. In general, there are five types of internal organisation : (1) the board of directors, (2) the chairperson, (3) the auditor, (4) some committees to assist the Council and (5) executive departments to carry out the detailed administrative matters.

“The Board of Directors” deals with important issues through active participation and discussion, and provides information such as major reports and publications from the council and research institutes. The board also assists its member institutes in making major policy by identifying their status and problems. It normally comprises a maximum of 12 directors, including the related central government ministers and vice ministers as well as professionals from academia, industry and the research institutes.

Regarding committees, there is “the Planning and Evaluation Committee” and “the Management Council”. The Committee coordinates research functions among member institutes, provides them with advisory services in relation to the long-term development vision and supports the evaluation of them. It normally comprises around 10 specialists from industry, academia and the research institutes. The Council provides advisory services on matters concerning member institutes such as improving the welfare of researchers. It also discusses how to reduce various administrative regulations and to maximise collaborative research among member institutes. It usually consists of members of the board of directors and presidents of member institutes.

As with the executive departments, there is a “General Secretary” comprising two teams, the “Planning and Management Team” and the “Evaluation Team” (KOCI, 2002c ; KORP, 2002 ; KRCP, 2002).

produce an evaluation manual with consultation from the OGPC (KOCI, 2002a).

## **(2) Evaluation Unit**

In this evaluation, the evaluation unit was “the Evaluation and Management Team” under the general secretary. As of April 2002, this team comprised five members of staff. Of these, three were directly engaged in the evaluation of member research institutes. As already mentioned, the mission of this team was to make a plan for the evaluation and to produce an evaluation manual after consultation with the OGPC and with assistance from the external support team (KOCI, 2002a).

However, in order to ensure the expertise of the evaluation exercise, the KOCI organised “a support team” on the basis of a contract, although the team was not formally responsible for the evaluation exercise. In the 2002 exercise, five researchers from the STEPI were approved as the support team. Their background was in management science, accounting, industrial engineering and so on (KOCI, 2002b : 4 ; KOCI, 2002c : i ; STEPI, 2002).

As a result, in the evaluation conducted by the KOCI in 2002, the evaluation and management team acted as the evaluation unit, even though it received much assistance from the support team.

## **(3) Evaluation Information**

The main source of information for the evaluation was “the performance report” submitted by each research institute. In general, this report is expected to contain the following information : an overview of the institute, including its mission and vision, the current organisational chart, the members of staff and the R&D budget for the year 2001, research outputs and management performance in 2001 and operating plans for 2002. The evaluation manual gave a detailed guideline for the report in relation to research outputs and management performance for the year 2001 (KOCI, 2002b ; 15-67). The major information requirements for the two evaluation objects are shown in <Table 5-16>

From these sources, it seems that both qualitative and quantitative information was adopted as the main source for evaluation, although the types and their weight varied slightly according to the main and sub evaluation objects.



<Table 5-16> Major Information Requirements for the 2002 Evaluation Exercise

Main Object	Sub Evaluation Object	Information Requirements
<b>R&amp;D Outputs (50%)</b>	Selection System for R&D Projects	<ul style="list-style-type: none"> <li>- state of projects in relation to institute's mission</li> <li>- procedure and criteria for project selection</li> <li>- application and success in project selection</li> <li>- list of R&amp;D projects conducted in 2001 etc.</li> </ul>
	Management System for R&D Projects	<ul style="list-style-type: none"> <li>- goal management and monitoring of projects</li> <li>- researchers involved in R&amp;D projects</li> <li>- system and procedure for evaluation etc.</li> </ul>
	Performance of R&D Project	<ul style="list-style-type: none"> <li>- good examples of indigenous R&amp;D projects</li> <li>- articles and proceedings published</li> <li>- IP including patent applied for and registered</li> <li>- successful examples of R&amp;D projects etc.</li> </ul>
	Specialisation of Research Field	<ul style="list-style-type: none"> <li>- R&amp;D structure in relation to institute's mission</li> <li>- efforts for specialisation in R&amp;D sector etc.</li> </ul>
	Collaborative Research	<ul style="list-style-type: none"> <li>- budget for collaborative research in 2001</li> <li>- list of collaborative researches among industry, academia and other research institutes</li> <li>- international collaborative research etc.</li> </ul>
	Exploitation and Transfer of Research Outputs	<ul style="list-style-type: none"> <li>- management system of research outputs</li> <li>- commercialisation and exploitation of research</li> <li>- technology transfer and technology export</li> <li>- royalties and spin-off companies</li> <li>- provision of public services and support etc.</li> </ul>
<b>Management Performance (50%)</b>	Establishment of Management Goals and its Achievement	<ul style="list-style-type: none"> <li>- achievement of major management goals</li> <li>- operating plans for 2002</li> <li>- mid and long-term development planning etc.</li> </ul>
	Internal Organisation and Humane Resources	<ul style="list-style-type: none"> <li>- flexibility of organisational management</li> <li>- current state of all staff</li> <li>- portion of supporting staff</li> <li>- evaluation system for researcher's performance</li> <li>- relations between evaluation and annual salary</li> <li>- women in science etc.</li> </ul>
	Management of Finance and Budget	<ul style="list-style-type: none"> <li>- sources of research fund</li> <li>- accumulation of retirement allowance etc.</li> </ul>
	Rationalisation of Management System	<ul style="list-style-type: none"> <li>- relationship with trade union</li> <li>- correction of findings of 2000 exercise</li> <li>- correction of audit results of 2001 etc.</li> </ul>

### 5.3.4.3. Evaluation Performance

#### (1) Evaluator

In order to secure the expertise and the objectivity of the evaluation, the OGPC and the KOCI have a policy to conduct the evaluation of the GRIs through the commission of evaluation exercise to outside experts or institutions (KOCI, 2002b : 4). In line with this policy, "the Committee for the Evaluation of Member Research Institutes" consisting of

17 external specialists conducted the KOCl's evaluation for 2002. As already noted, this committee comprised members of both the planning and evaluation committee of the KOCl, and other specialists who came from outside in the areas of knowledge management, accounting, human resources and organisation theory. Initially, the planning and evaluation committee comprised experts from academia, the research institutes and industry, whose major interest related to S&T areas of individual research institutes, such as oriental medicine, industrial technology, IT technology & computer science, food science, machinery & materials, chemistry and electrotechnology. The committee was again divided into two sub committees in accordance with the two evaluation objects : the sub committee for R&D matters and the sub committee for internal management system. (KOCl, 2002c : 12-13). The detailed membership of the evaluation committee is shown in <Table 5-17>.

**<Table 5-17> Membership of Evaluation Committee in the KOCl's 2002 Exercise**

<b>Institution Sub Committee</b>	<b>Academia</b>	<b>Industry</b>	<b>Research Institute</b>	<b>Total</b>
<b>R&amp;D Sub Committee</b>	5	3	2	10
<b>Management Sub Committee</b>	4	1	2	7
<b>Total</b>	9	5	4	17

Therefore, the evaluation committee took actual responsibility for the evaluation of the seven member research institutes in 2002. It conducted the evaluation through a literature review and a site visit, and then produced the evaluation report and made some useful recommendations.

## **(2) Evaluation Method**

The committee adopted qualitative evaluation methods including a literature review and a site visit. The committee firstly reviewed the performance report submitted by the individual research institutes in advance of the exercise. Secondly, it visited each institute for two and a half hours, in order to confirm the contents of the performance report. On these visits, it conducted interviews with some related members of staff to get supplementary information (KOCl, 2002b : 14 ; KOCl, 2002c : 7).

Consequently, it appears that, in the evaluation of research institutes under the KOCI, the key evaluation methods adopted were a literature review and a site visit including interviews, which are regarded as typical types of qualitative methods.

In addition, the KOCI conducted a survey to confirm the customers' satisfaction with each institute, even though the results were indirectly reflected in the final evaluation results. This survey covered the following six matters for each institute : (1) the performance and the possibility of development, (2) the management of the organisation and research resources, (3) the quality of research results, (4) collaborative research with industry, academia and other research institutes, (5) the level of provision of related services and technology and (6) the relationships and continuity with related customers. A total of 1,300 questionnaires were distributed to government ministries, other GRIs, research institutes in the public sector, universities, large-scale industry, small and medium size enterprises (SMEs) and others who had been in collaboration or cooperation with member institutes under the KOCI. Around 550 were returned. These were analysed in both a quantitative and qualitative way and also utilised as a source of reference information for the evaluation exercise (KOCI, 2002c : 53-74).

### **(3) Evaluation Criteria and Indicator**

In order to judge each sub evaluation object, one to four evaluation criteria that were expected to reflect the attributes of each evaluation object were adopted. Then, each evaluation criterion was again divided into specific evaluation indicators. As a result, in the evaluation of research institutes under the KOCI in 2002, a total of 22 evaluation criteria and 45 evaluation indicators were adopted (KOCI, 2001c : 9-11). The detailed evaluation criteria and the number of indicators adopted in relation to each criterion are shown in <Table 5-18>.

For individual indicators, the evaluation was conducted firstly in a qualitative way and then, in principle, a certain grade was given on a five-tier grading system, though a three or a ten-tier grading system was also used, where appropriate. Meanwhile, for the indicators with a three or a five-tier grading system, the committee adopted an absolute evaluation method to judge the performance of individual research institutes independently without then comparing others. On the other hand, in the case of the indicators with a ten-tier grading system, it adopted a relative evaluation method to

investigate each institute's performance in comparison with others (KOCI, 2002c : 8).

**<Table 5-18> Evaluation Criteria and Indicators for the GRIs under the KOCI**

<b>Main Object</b>	<b>Sub Evaluation Object</b>	<b>Evaluation Criteria</b>	<b>Evaluation Indicators</b>
<b>R&amp;D Outputs (50%)</b>	Selection System for R&D Projects	- appropriateness of research planning - rationality of appraisal process - allocation state of R&D projects	7
	Management System for R&D Projects	- appropriateness of R&D management system - rationality of output evaluation system	3
	Performance of R&D Project	- performance of indigenous projects - quality of academic activities - quality of intellectual properties - flagship research outputs	9
	Specialisation of Research Field	- degree of specialisation and concentration in research field in relation to missions	2
	Collaborative Research	- collaborative research among industry, academia and other institute - international collaborative work	3
	Exploitation and Transfer of Research Outputs	- systematisation of output management - exploitation of research outputs - provision of public services	5
<b>Management Performance (50%)</b>	Establishment of Management Goals and its Achievement	- achievement of management goals - appropriateness of operating planning - rationality of development plans	4
	Internal Organisation and Human Resources	- appropriateness of organisational management - rationality of HR management - evaluation system for researcher's performance	5
	Management of Finance and Budget	- soundness of financial structure	3
	Rationalisation of Management System	- rationalisation of internal management - correction of findings in 2001 - credibility of performance report	4

Consequently, in the KOCI evaluation, it is apparent that evaluation criteria and indicators adopted were generally qualitative, although a few quantitative ones were utilised.

#### **5.3.4.4. Evaluation Utilisation**

##### **(1) Evaluation Report**

The evaluation report was published in April 2002 and is available on the KOCI website

(<http://www.koci.re.kr>). The report was produced by the evaluation committee, with administrative assistance from the support team. As in other cases, this report covers evaluation components including an evaluation purpose, evaluation objects, evaluation methods, the list of evaluators, the evaluation results of member research institutes, the results of the survey of customers' satisfaction, some recommendations and so on. In particular, in the section concerning the evaluation results of individual research institutes, it deals with the synthesis of the evaluation findings for each main evaluation object and recommendations for future development (KOCI, 2001c).

## **(2) Utilisation Type**

In principle, the findings of the evaluation of member institutes under the research councils can be utilised, where appropriate, to close any institute that shows bad progress, as well as to dismiss the president of any institute who is considerably lacking in management ability and the capability necessary for his position. In addition, each institute has to reflect the evaluation findings in the strategies and plans of the institute management and research programmes for the following year (OPM, 2001c).

In accordance with this policy, the evaluation findings of the KOCI member institutes were submitted to the Prime Minister and the Minister of the MPB. Each member institute was required to reflect these findings in their 2002 strategies and plans for internal management and R&D programmes. The correction work suggested by the evaluation findings will also be regularly monitored by the KOCI (KOCI, 2001c : 3, 80). Hence, regarding the evaluation utilisation, it seems that more emphasis was placed on the short-term improvement of the internal management system and R&D management at the level of each research institute, rather than on the mid and long-term development of R&D activities in the public sector at the national level.

### ***5.3.5. Evaluation of R&D Project : Evaluation of the Creative Research Initiatives (CRI) by the MOST (2000)***

The MOST supports a number of research areas through specialised initiatives or inter-departmental programmes where the need for government investment for the improvement of national competitiveness and the quality of life has been identified. The

research programmes of the MOST cover areas from basic science to long-term strategic R&D programme. It is the MOST's policy to evaluate all of its research projects and to use the findings to inform future funding decisions. So, on the basis of the evaluation findings, it is expected that some lower ranked projects will be cut. According to this policy, between June and September 2000, the MOST, through evaluation exercise entrusted to the KISTEP, evaluated the interim achievements of 27 R&D projects under the Creative Research Initiatives (CRI) that had been started in 1997, which was one of several sub R&D programmes of the MOST. It was the first time that the MOST had conducted the evaluation of R&D projects under the CRI.

#### ***5.3.5.1. Outline of the Creative Research Initiatives (CRI) of the MOST***

The launching of the Creative Research Initiatives in 1997 symbolised a policy shift in S&T development in Korea from imitative to creative innovation, and towards the knowledge-based economy. The aim of this programme is to strengthen the national potential for technology competitiveness through creative basic research.<sup>12</sup> The CRI focuses on exploring various phenomena that will happen in the physical environment, developing new fields of scientific research, and making technological breakthroughs. The CRI research funds support areas of research based on creativity and originality. So, the CRI emphasises a high level of flexibility in research in order to enhance creativity. The project leader, who is selected by strict criteria based on creativity, leadership, research experience and so on, has exclusive authority and responsibility in managing the project (MOST, 2002c).

The funding areas are as follows : (1) the scientific investigation of natural and artificial phenomena, (2) the creation of new research fields through the scientific investigation of natural and artificial phenomena, including bio-mechanisms, reactions on fields, photosynthesis and so on, (3) exploration of emerging research fields in science and technology, (4) exploration and development of revolutionary ideas emerging from new research fields, including AI, bionic, molecular robots, smart materials and so on, (5) the

---

<sup>12</sup> In order to achieve the policy objectives of the CRI set up initially, the MOST adopts relatively advanced approach of evaluation for this programme, very different from those for other programmes. These include the introduction of, a so-called, open evaluation system through the active involvement of international experts and volunteer reviewers as well as an emphasis on the evaluation the quality of core research outputs through the simplification and quantification of evaluation indicators (MOST and KISTEP,

achievement of breakthroughs in conventional technology areas and (6) research geared towards overcoming the limitations of existing science and technology, including optics, high efficiency energy conversion, environmental protection or recovery, etc. (Lee, 2002a : 40). Each project is normally funded to the extent of three hundred million KRW to 7.5 hundred million KRW per year, depending on its attributes. Currently, a total of 57 projects have been supported and a total of 143.3 billion KRW has been invested on these projects (MOST, 2002h). The detailed numbers of projects and investments of the CRI in each year are shown in <Table 5-19>.

**<Table 5-19> Project Numbers and Fund Investment for the CRI (1997 to 2001)**

	Project Number		Fund Investment (hundred million KRW)				
	Selected	Current	1997	1998	1999	2000	2001
1997	27	22	200	197	167	113	110
1998	19	15	-	142	117	88.2	79
1999	6	6	-	-	37	28.8	30
2000	8	8	-	-	-	48	40
2001	7	6	-	-	-	-	36
<b>Total</b>	67	57	200	339	321	278	259

- Note : 1. Of the 27 projects in 1997, five ended after the first stage evaluation in 2000.  
 2. Of the 19 projects in 1998, one stopped because of the project leader's death and three projects ended after the first stage evaluation in 2001.  
 3. In the year 2001, one project was cancelled because project leader left his position.

Meanwhile, the Korean government plans to conduct a general evaluation of the CRI in 2003, after six years of its implementation. The future direction of the CRI will be decided upon after this evaluation (MOST, 2002c).

### **5.3.5.2. Evaluation Paradigm**

#### **(1) Evaluation Purpose**

This evaluation was conducted in order to decide whether 27 R&D projects started in the 1997 would be given further research funding, through the investigation of final and interim research outputs. The MOST conducted this evaluation to achieve the policy

---

2000a : 1). It was for this reason that I adopted the evaluation exercise of the CRI as a case study.

objectives of the programme through the judgement of the quality of research as well as to maintain the original concepts and framework of the programme through the appropriate response to the trends of technological change (MOST and KISTEP, 2000b : 1).

In case of the evaluation of the CRI's projects, the evaluation purpose indicated above represented the terms of reference. So, it seems that this exercise also concentrated on the "accountability" of the project leader and the "improvement" of the programme itself, like the evaluation of other R&D projects.

## (2) Evaluation Object

As already indicated, as at June 2000, this exercise included 27 R&D projects started in the year of 1997. These projects were categorised into four research areas and carried out by universities, the research institutes and industry (MOST and KISTEP, 2000b : 2, 5). The detailed research areas and types of performers are shown in <Table 5-20>.

**<Table 5-20> Research Areas and Performers for 1997 Projects under the CRI**

Research Areas				
Chemistry & New Compounds	Life Sciences & Medicine	Physics and Materials	Mechanical E. & Electronics	Total
7	10	5	5	27
Research Performer				
Universities	Research Institutes	Industry	Total	
24	1	2	27	

Therefore, it appears that the level of the evaluation object was the individual projects that had made up of the CRI programme. As already indicated, as at June 2000, although there were a total of 60 projects under the CRI programme that had been selected in 1997, 1998, 1999 and 2000, the 2000 evaluation exercise covered only R&D projects implemented since 1997 (MOST, 2002b). Thus, the scope of the evaluation object was confined to existing projects, which had already produced some research outputs in the past three years.

Meanwhile, each R&D project was evaluated from the perspective of four evaluation



criteria : (1) the achievement of research objectives, (2) originality of research outputs, (3) potential as a world-class researcher for the next generation and (4) the appropriateness of future research plans (MOST and KISTEP, 2000a : 12).

### **(3) Evaluation Type**

This evaluation mainly aimed at examining and judging what individual R&D projects had achieved in terms of research outputs and their quality, the capability of the project leader, future research plans and so on (MOST and KISTEP, 2000a : 12). It is apparent that this evaluation focused on an “output” evaluation and an “assessment” of future research strategy, which were recognised as typical types of R&D assessments.

#### **5.3.5.3. Evaluation Resources**

##### **(1) Evaluation Manpower**

Two groups of evaluation manpower were engaged in this evaluation. First of all, there were many external experts such as (1) evaluation panels, (2) international evaluators, (3) volunteer reviewers through the internet and (4) the planning and management committee (PMC) for the CRI. The others were the internal members of staff of “the Division of National R&D Planning and Management”, one of the five divisions in the KISTEP.

Firstly, for this evaluation, “four sub evaluation panels” were organised, one for each of the four research fields. They had a principal mission to conduct the evaluation through the incorporation of results from the international evaluators and volunteer reviewers. In general, each sub evaluation panel comprised around 12 members, who were the members of the PMC for the CRI, related programme coordinators of the KISTEP responsible for overall R&D management in individual research area <sup>13</sup> or other professionals randomly selected from the evaluator pool, which consisted of many

---

<sup>13</sup> In the Division of National R&D Planning and Management, there are seven programme coordinators (PCs) to manage overall R&D activities including the survey of technological trends, research planning, appraisal, evaluation and so on. These are (1) the PC for mechanical engineering, (2) the PC for information technology and electronics, (3) the PC for health and life sciences, (4) the PC for materials and chemistry, (5) the PC for energy and environmental technology, (6) the PC for space science and fundamental research and (7) the PC for nuclear energy technology. They are normally seconded to the KISTEP from related GRIs and universities for a period of two years. In addition, in principle, each coordinator is given the responsibility for overall management of one or two sub national R&D programmes supported by the MOST. The Programme Coordinator for Space Science and Fundamental Research is in charge of the CRI programme (KISTEP, 2002b).

experts from industry, academia, the research institutes and so on. A total of 50 external experts were involved in the 2000 exercise as members of four evaluation panels (MOST and KISTEP, 2000b : 3).

Secondly, in an effort to secure the professionalism and fairness of the evaluation, prominent international experts also played a major role in the evaluation. In principle, there were three “international evaluators” for each project including one expert recommended by the pertinent project leader. In the case of the 2000 evaluation, a total of 77 international experts from 12 countries evaluated the 26 R&D projects (MOST and KISTEP, 2000b : 3).

Thirdly, some “volunteer reviewers” were invited to the evaluation through the internet. In general, a maximum of 30 reviewers are allowed for each project and they are required to have carried out the research concerned for over 10 years. In the 2000 evaluation exercise, the KISTEP encouraged related researchers either in Korea or overseas to join the evaluation work through the announcement on the website of 19 domestic learned societies including the Korean Physical Society (KPS) as well as of 7 overseas Korean scientists and engineers associations including the Korean-American Scientists and Engineers Association (KSEA). The evaluation results of volunteer reviewers were utilised as important reference information for the evaluation panels (MOST and KISTEP, 2000a : 9 ; MOST and KISTEP, 2000b : 6).

Finally, “the Planning and Management Committee for the CRI” was involved in the evaluation through the review and synthesis of the results from other evaluators. This committee was initially organised to set up the basic strategy for the CRI programme and then fully engaged in the overall management of that programme, including the selection of new projects, an annual monitoring and an output evaluation. The members of the committee were a total of 18 who comprised 16 prominent scientist from industry, academia and the research institutes as well as two related staff in the MOST and the KISTEP (MOST and KISTEP, 2000b : 7, 12-13). The detailed membership of the PMC for the CRI is shown in <Table 5-21>.

In addition, six members of staff from “the PC for Space Science and Fundamental Research”, one of the seven PCs in the Division of National R&D Planning and

Management, were partially involved in the evaluation through their preparation of evaluation plan and management of the detailed evaluation process (KISTEP, 2002b).

**<Table 5-21> Detailed Membership of the PMC for the CRI**

<b>Institution</b>	<b>Academia</b>	<b>Industry</b>	<b>Research Institutes</b>	<b>MOST (Director)</b>	<b>KISTEP (PC)</b>	<b>Total</b>
<b>Number</b>	7 (39%)	4 (22%)	5 (28%)	1 (5.5%)	1 (5.5%)	18

As a result, around 140 external experts from both Korea and overseas were directly engaged in the evaluation of 27 R&D projects under the CRI, while many scientists were also indirectly involved in this exercise through conducting the evaluation on the internet.

## **(2) Evaluation Unit**

As already described, the Programme Coordinator for Space Science and Fundamental Research is generally responsible for the overall management of the CRI programme. In line with this role, the coordinator prepared the evaluation plan and managed the whole of the detailed evaluation process. (KISTEP, 2002b). Thus, it is apparent that this coordinator acted as the evaluation unit in the evaluation for the CRI programme of the year 2000, even though he received a great deal of assistance from other PCs concerned. As at October 2002, there were six member of staff under the PC for Space Science and Fundamental Research.

## **(3) Evaluation Information**

In advance of the evaluation, each project leader was required to submit materials necessary for the exercise. Unlike in other Korean evaluation exercises, it was expected that all evaluation materials would be written in English (MOST and KISTEP, 2002a : 5-6, 10). The detailed information requirements for each evaluation criteria are shown in <Table 5-22>.

Therefore, in the evaluation of R&D projects under the CRI programme, it appears that while qualitative and quantitative information were utilised equally, quantitative information was considered more important than in other Korean evaluation exercises mentioned above, in that most information regarding both the originality of the research outputs and potential as a world-class research leader can be counted and measured.

<Table 5-22> Information Requirements for the 2000 Exercise for the CRI

Evaluation Criteria	Information Requirements
<b>Achievement of Research Objectives (20)</b>	- research achievements comparing to original objectives
<b>Originality of Research Outputs (40)</b>	- articles published in SCI listed journals and their citation record (in principle, articles in SCI journals with an impact factor over five) - three excellent published articles and opinions of their peer reviewers - patents applied for and registered
<b>Potentials as a World-class Research Leader (20)</b>	- invitations to international conferences etc. - presentations at international seminars etc. - prizes and other measures of esteem
<b>Appropriateness of Future Research Plans (20)</b>	- research plan and strategy for the second stage

\* *Figures in parenthesis represent the weight of each evaluation criterion*

#### 5.3.5.4. Evaluation Performance

##### (1) Evaluator

As already mentioned, four types of external expert or expert group, namely an evaluation panel, international evaluators, volunteer reviewers and the PMC for the CRI, conducted the evaluation of the R&D projects selected in 1997. However, there were slight differences in their roles, in that while the results produced by both the evaluation panel and international evaluators were reflected in the final evaluation, those produced by the volunteer reviewers were only utilised as reference information for the exercise conducted by evaluation panels. Also, unlike the evaluation panel, neither the international evaluators nor the volunteer reviewers conducted an evaluation of the achievement of objectives, which was one of the four evaluation criteria. The results of each evaluator group were reviewed and synthesised by the PMC for the CRI (MOST and KISTEP, 2000a : 5 ; MOST and KISTEP, 2000b : 4, 7).

As a result, it appears that the roles of the evaluation panel and international evaluators were more important than those of other two types of evaluator, even though they all related to the evaluation of the 27 R&D projects under the CRI. For this reason, further discussion will relate only to the evaluation panel and the international evaluators. The

detailed roles of each evaluator group are summarised in <Table 5-23>.

**<Table 5-23> The Role of the Evaluator Groups in the 2000 Exercise of the CRI**

<b>Evaluator Evaluation Criteria</b>	<b>Evaluation Panels</b>	<b>International Evaluators</b>	<b>Volunteer Reviewers</b>	<b>PMC for The CRI</b>
Achievement of Research Objectives	Yes	No	No	- review and synthesis of results of other evaluators - production of final evaluation results
Originality of Research Outputs	Yes	Yes	Yes	
Potential as a World-class Researcher	Yes	Yes	Yes	
Appropriateness of Future Research Plans	Yes	Yes	Yes	
<b>Utilisation for Final Findings</b>	<b>50% reflected</b>	<b>50% reflected</b>	<b>Reference Information</b>	

The evaluation panel was again divided into four sub evaluation panels one for each research field. As already noted, each sub evaluation panel comprised around 12 members including the members of the Planning and Management Committee (PMC) for the CRI, related programme coordinators of the KISTEP and other professionals from industry, academia and the research institutes (MOST and KISTEP, 2000b : 5). The detailed membership of the evaluation panel is shown in <Table 5-24>.

**<Table 5-24> Membership of Evaluation Panel for the CRI 2000 Exercise**

<b>Institution</b>	<b>Academia</b>	<b>Industry</b>	<b>Research Institutes</b>	<b>Total</b>
<b>Number</b>	23 (46%)	14 (28%)	13 (26%)	50

In addition, some international experts were involved in carrying out the evaluation of R&D projects under the CRI. Each project was evaluated by three prominent international experts from the research areas concerned. They were drawn from evaluator pool, which was recommended by the pertinent project leader, related programme coordinators and the PMC for the CRI. In the case of the 2000 exercise, a total of 77 international experts were involved in the evaluation (MOST and KISTEP, 2000b : 6). The details of their nationality and the institutions they belong to are summarised in <Table 5-25>.

<Table 5-25> Details of International Evaluators in 2000 CRI Exercise

Nationality			
Country	Number	Country	Number
USA	51 (66%)	Belgium	2
Japan	9 (12%)	Italy	1 (1%)
UK	3 (4%)	Hungary	1
Switzerland	3	Canada	1
Germany	2 (3%)	Taiwan	1
France	2	Austria	1
Institution They Belong to			
Academia	Industry	Research Institute	Total
55 (72%)	7 (9%)	15 (19%)	77

Consequently, the external evaluators played a principal role in the 2000 evaluation of the CRI. Among the external evaluators, both the member of the evaluation panels and international experts conducted the actual evaluation, although volunteer reviewers and the PMC for the CRI were also engaged in producing useful information for it.

## (2) Evaluation Method

The principal evaluation methods adopted were the analysis of the evaluation materials submitted by the project leader and the bibliometric analysis of the research outputs. In addition, a site visit was also undertaken, if necessary.

Both the sub evaluation panels and international experts evaluated the excellence and the originality of research outputs of each project, on the basis of the qualitative analysis of the evaluation information indicated above. In particular, the evaluation panel conducted an evaluation of the overall achievements of each project. In each such evaluation, the project leader gave an open presentation about overall performance of the project (MOST and KISTEP, 2000b : 4).

A bibliometric analysis of the research outputs in terms of articles and patents was also conducted for this evaluation. In particular, calculations were based on research articles published in Science Citation Index (SCI) listed journals. A greater weight was given to articles published in SCI journals ranked with an impact factor over five (MOST and KISTEP, 2000a : 6).

Meanwhile, for each evaluation criterion, both the evaluation panel and international evaluators were required to give a certain grade on a four-tier grading system of excellent, very good, good and fair (MOST and KISTEP, 2002b ; 5-6).

As a result, although qualitative and quantitative evaluation methods were both adopted, greater emphasis was placed on quantitative methods and, in particular, on the bibliometrics of research articles than in other evaluation exercises. It seems that this resulted from the principle of the CRI evaluation exercise, the intention of which was to judge the quality of core research outputs using the more easily quantified evaluation criteria and indicators possible.

### (3) Evaluation Criteria and Indicators

As already mentioned above, in the evaluation system of the CRI programme, the MOST and KISTEP have as a principle to ensure the expertise and fairness through the active involvement of international experts and volunteer reviewers as well as to evaluate the quality of core research outputs through the simplification and quantification of evaluation indicators (MOST and KISTEP, 2000a : 1).

**<Table 5-26> Evaluation Criteria and Indicators in 2000 CRI Exercise**

<b>Evaluation Criteria</b>	<b>Weight</b>	<b>Evaluation Indicator</b>
Achievement of Research Objectives	20%	None
Originality of Research Outputs	40%	- quality of articles published
		- opinions of peer reviewers on articles submitted to the SCI journals
		- quality of patents applied for and registered
Potential as a World-class Researcher	20%	- international academic activities
		- academic reputation both home and overseas
Appropriateness of Future Research Plans	20%	None
4 Criteria	100%	5 Indicators

In accordance with this principle, in the 2000 exercise, four evaluation criteria and five evaluation indicators were adopted, relatively fewer than in other cases. Meanwhile, each evaluation criterion was given a different weight, depending on its attributes (MOST and

KITEP, 2000a : 5 ; MOST and KISTEP, 2000b : 5). The details of the evaluation criteria and indicators are shown in <Table 5-26>.

#### **5.3.5.5. Evaluation Utilisation**

##### **(1) Evaluation Report**

The evaluation report was prepared by the KISTEP. This report contained the general information including the list of projects, the evaluation purpose, the evaluation process, evaluation methods, the PMC's integrated evaluation results for the 1997 selected R&D projects and so on (MOST and KISTEP, 2000b).

This report was internally distributed to the staff of the MOST and the KISTEP as well as to members of the evaluation panel and the PMC for the CRI. However, it was not made available to the public, either electronically on the MOST and the KISTEP website or in hard copy form.

##### **(2) Utilisation Type**

In general, the findings of the MOST evaluation exercise have been utilised to decide whether R&D projects should be awarded further research funds. As a part of this policy, it is expected that around 15 % of the lower ranked projects in the evaluation of the CRI programme will be cut (MOST, 2002f ; MOST, 2002g). In line with the general policy of the MOST, the evaluation results for the projects under the CRI programme have also been used to improve the future structure of the programme.

The comprehensive results and final grades for each project arrived at by the PMC for the CRI summarised in <Table 5-27>. On the basis of these findings, of the 27 R&D projects, five that have a grading of either "fair" or "unsatisfactory" were terminated, whereas the rest of them continued to get research fund (MOST and KISTEP, 2000b : 12-13).

**<Table 5-27> Evaluation Results in the 2000 CRI Exercise**

<b>Grade</b>	<b>Very Good</b>	<b>Good</b>	<b>Fair</b>	<b>Unsatisfactory</b>	<b>Total</b>
<b>Number of Projects</b>	7	15	4	1	27
<b>Proportion</b>	26.0 %	55.5 %	14.8 %	3.7 %	100 %



### ***5.3.6. Synthesis and Conclusion***

In this section, case studies have been conducted of five evaluation exercises in the S&T and R&D sectors in Korea. The synthesis of these case studies is given in <Table 5-28>, on the basis of the analyses of the individual evaluation exercises.

From this synthesis, it appears that, in Korea, R&D evaluation exercises, except in the case of R&D project evaluation of the MOST, have started to be carried out relatively in recent, with most evaluation system being introduced in the late 1990s and early 2000s. As a result, it seems that, generally in Korea, evaluation exercises in these sectors have similar characteristics to most aspects of meta-evaluation components, including evaluators and evaluation utilisation. All five evaluation exercises were conducted by independent evaluation committees, which comprised external experts from industry, academia, the research institutes and the media. So, it appears that external evaluators had an actual responsibility for the exercise, even though they received much assistance from internal staff in related institutions. Budget adjustment or allocation was considered as the most frequent and important use to which the evaluation findings were put. Meanwhile, another feature of the Korean R&D evaluation system is that all evaluation exercises have been performed every year without consideration of the individual attributes and characteristics of each evaluation system, even though there may be a great deal of difference between the evaluation of S&T and R&D policy, the upper level of the national R&D evaluation hierarchy shown in (Figure 5-2), and the evaluation of R&D projects under the CRI programme, the lowest level of the national R&D evaluation system.

However, it seems that, in contrast to the other cases, the evaluation exercise of R&D projects under the CRI programme represents a recent policy change undertaken by the Korean government to secure the objectiveness and professionalism of the evaluation exercise concerned. As a part of this policy, all R&D projects under the CRI programme are required to be evaluated by prominent international experts. Also, a great deal of emphasis was placed on research articles that had been published in SCI listed journals, in particular, in those with an impact factor over five.

<Table 5-28> Synthesis of the Evaluation Exercises in S&T and R&D Sector of Korea

Evalu. Components		Cases	Major Policy Issues of the MOST(2001)	National R&D Programmes (2001)	Five Research Councils (2001)	GRI's under the KOCI (2002)	R&D Projects of the CRI (2000)
Evaluation Paradigm	Purpose		- accountability - policy/programme improvement	- accountability - programme improve. - knowledge sharing	- accountability - improvement of RCs' management - RC's own activities - support of their GRI's - internal management	- management improve. - accountability - knowledge transfer	- accountability - decision for future fund - programme improve.
	Object		- sub programme (4) - main project (17)	- sub programme (161) - sub project (9,533)	- output evaluation - assessment of R&D plan	- R&D matters - internal management	- R&D project started in 1997
	Type		- interim evaluation - output evaluation - impact evalu. (optional)	- interim evaluation - output evaluation	- evaluation comm. (10) - DG for R&D support under OGPC (6) - National R&D Evalu. Div.in KISTEP (2)	- output evaluation - prospective assessment	- output evaluation - assessment of future research plan
Evaluation Resources	Manpower		- PEC (1) - support team (2) - AMO for PAE (2)	- steering committee (9) - programme comm. (83) - review committee (140) - supporting team (16) - MOST(8) / KISTEP(21)	- evaluation comm. (10) - DG for R&D support under OGPC (6) - National R&D Evalu. Div.in KISTEP (2)	- evaluation commi. (17) - support team (5) - Evalu. & Management team in KOCI (3)	- evaluation panel (50) - overseas experts (77) - volunteer reviewers - PMC for CRI (16) - staff of PC for SS & FR (6)
	Unit		- Director-General I for PAE under OGPC	- STA Division in MOST - National R&D Evalu. Division in KISTEP	- DG for R&D support under OGPC • 2 teams involved	- Evalu. & Management team in KOCI with outside support team	- PC for SS & FR in KISTEP
	Information Sources		- self-evaluation report - operating plan - extensive consultation	- performance report - presentation by programme manager	- performance report - questionnaire survey	- performance report	- research abstract - articles & patents
Evaluation Performance	Evaluator		- external > internal	- external	- external	- external	- external
	Method		- literature review - in-depth interviews	- literature review	- literature review - site visit - questionnaire survey	- literature review - site visit - in-depth interview	- review of materials - bibliometric analysis
	Criteria/ Indicator		- qualitative criteria/items • 6 criteria, 10 items - no clear identification of evaluation indicators	- qualitative criteria • programme evau. (6) • project review (2)	- qualita. > quantitative • 21 criteria • 36 indicators	- both qualita. and quanti. • 22 criteria • 45 indicators	- quantitative indicator • articles & patents - qualitative criteria • goal achievement etc.
Evaluation Utilisation	Report		- published - submitted to the NA - posted on website	- published - posted on website • NSIC, MOST	- published - not available on website	- published - posted on website	- not open to the public - internal circulation
	Utilisation Type		- budget adjustment - corrective measures • 5 projects identified	- budget adjustment - improvement measures • program (70) : 7 types • project (164) : 2 types	- budget adjustment - coordinating national R&D programmes of GRI's	- budget allocation - reflected in strategies and plans for 2002	- projects with bad evaluation were cut

## **CHAPTER 6. META-EVALUATION AND COMPARSION OF THE UK AND THE KOREAN CASES**

### **6. 1. Introduction**

In Chapters 4 and 5, I conducted case studies of the UK and the Korean national R&D evaluation system. These included five evaluation exercises : the evaluation of S&T and R&D policy, the evaluation of S&T or R&D programmes, the evaluation of research councils, the evaluation of the research institutes under the responsibility of the research councils and the evaluation of R&D projects in each country.

As a result, it appears that, in the UK system, in general, the five evaluation exercises have similar features, even though there are slight differences in certain meta-evaluation components such as evaluation resources and evaluation performance. On the other hand, there are also many similarities in the Korean R&D evaluation exercises in most aspects of meta-evaluation components, although the evaluation of R&D projects under the CRI programme differs from that of other cases, in that extra efforts, including the active involvement of international evaluators, have been made to secure the fairness and professionalism of the evaluation exercise. In addition, in both the UK and the Korean cases, it is clear that the evaluation findings of all evaluation exercises have been fully or partially reflected in decision-making processes regarding future budget allocation and further research funding.

This chapter comprises three main sections excluding introduction : earlier two sections are meta-evaluations of the UK and the Korean R&D evaluation exercises recently completed, while the last one is a comparison between the national R&D evaluation systems in the UK and Korea as a whole. Firstly, on the basis of the results of the two case studies, I have carried out the meta-evaluation of the UK and the Korean national R&D evaluation systems, using the meta-evaluation model consisting of four main component and 11 sub components suggested in Chapter 3. Through conducting the meta-evaluations, the individual strengths and limitations of the UK national R&D evaluation system and

the Korean system can be identified. The findings of the meta-evaluation will be utilised as referential information for the comparison of the two nations' national R&D evaluation systems.

Then, the UK and Korean national R&D evaluation systems will be compared. As already indicated, in principle, this comparison is carried out on the basis of the findings of the meta-evaluation exercises conducted in two previous sections. In addition, in order to provide a more comprehensive understanding of the two national R&D evaluation systems, the hierarchy of the entire R&D evaluation systems at the national level and the evaluation timing are used as useful extra information for the comparison, even though these are excluded from the meta-evaluation components. Through this comparison, I shall identify whether there is any possibility of applying the elements of the UK R&D evaluation system to the Korean case, and to investigate the extent to which the UK system and its practice could be applied to the Korean case. Therefore, these results, as well as the findings of the meta-evaluation, will be utilised as the basic materials upon which suggestions for the applicability of the UK national R&D evaluation system to the Korean system are based.

## **6. 2. Meta-evaluation of the UK National R&D Evaluation System**

In this section, using the meta-evaluation model suggested in <Table 3-4>, I shall analyse and judge the effectiveness, rationality and appropriateness of the five R&D evaluation systems and their exercises in the UK. In particular, I shall examine the extent to which the current state of the individual R&D evaluation systems approaches the ideal state of the evaluation system from the perspective of policy evaluation theory. As ready noted, the results of the five case studies of R&D evaluation systems carried out at the different level in Chapter 4 were utilised, in order to conduct the meta-evaluation.

Meanwhile, unlike the case studies that were conducted separately under independent and individual sub titles relating to the 11 sub meta-evaluation components, this meta-evaluation will be carried out under more integrated titles on the basis of the four main components of the meta-evaluation model. This is to avoid the repetition of some related contents and information between the case studies and the meta-evaluations.

### ***6.2.1. Meta-evaluation of the Cross-cutting Review of Science and Research in 2000***

The 2000 cross-cutting review of science and research was undertaken in order for many government departments concerned with the S&T policies and programmes to develop co-operation with one another and to make a variety of new working arrangements. In the light of the principal policy objectives, I judged and examined the rationality of the evaluation system and the appropriateness of the actual practice of the 2000 cross-cutting review of science and research conducted as a part of the 2000 Spending Review (SR).

#### ***6.2.1.1. Evaluation Paradigm***

As already discussed, the evaluation paradigm, consisting of (1) the evaluation purposes, (2) the evaluation objects and (3) an evaluation type, fundamentally prescribes the basic directions, content and characteristics of the overall evaluation activities. This is because these components determine the input of evaluation resources, the substance of evaluation performance and the types of evaluation utilisation that are performed after setting up the evaluation paradigm. Therefore, their rationality and appropriateness must firstly be examined and judged in the meta-evaluation.

(1) Evaluation Purpose : In this review, the evaluation purpose can be drawn from the remit of the 2000 cross-cutting review of science and research (HM Treasury, 2000a : 2). As a result, as in most policy evaluations, both the “accountability” of related government departments and organisations and the “improvement” of policies and programmes concerned were announced explicitly as two main evaluation purposes. Regarding the evaluation purposes, most researchers suggest “accountability”, “improvement of policy and programme” and “enhancement of basic knowledge” as the rational purposes of policy evaluation (Goldenberg, 1983 ; Arvidsson, 1986 : 627 ; Lee & Sampson, 1990 : 157 ; Hudson, Mayne and Thomlison, 1992 : 5), although there are some discussions about less rational purposes of policy evaluation including the posture of policy effects, the postponement of decision making, the ducking of responsibility, public relations and the eyewash of policy effects (Weiss, 1972 : 11-12 ; Suchman, 1972 : 81).

From the above discussions, it seems that the two evaluation purposes of the 2000 cross-cutting review of science and research are rationally established.

However, in addition to the two evaluation purposes, it is necessary to adopt “the enhancement of basic knowledge” as the additional evaluation purpose of the cross-cutting review of science and research, although it is not clearly revealed in the official documents concerned. This type of evaluation purpose will help policy makers and related researchers to confirm the existing theories as well as to develop new approaches in the areas of S&T and R&D policy.

(2) Evaluation Object : The 2000 cross-cutting review examined the science and research activities supported by public funds through the Higher Education Funding Councils (HEFCs) and research councils as well as conducted by universities and public sector research establishments (PSREs) funded by other civil government departments (HM Treasury, 2000a : 2). We know that the evaluation object of this review included both science and research “policies” in civil departments and their main S&T and R&D “programmes”. Therefore, it is apparent that the setting up of the evaluation objects was rational in that this review included both the related government policies as well as their main programmes. This is because a certain policy normally comprises some detailed programmes so that it can also be directly affected by the overall achievements of these.

In addition, another feature of the evaluation object is that this review included policies and programmes of “all central government departments” and “their main Non-Departmental Public Bodies (NDPBs)” including the HEFCs and the research councils which are mainly responsible for SET activities. Thus, it is very proper that the evaluation object of the 2000 cross-cutting review of science and research extended to the S&T related all public organisations, rather than being confined to some departments. This is because there have recently been fast increasing interdisciplinary and multidisciplinary trends in the S&T and R&D sectors.

(3) Evaluation Type : Regarding the evaluation type, it appears, from the analysis of its remits, that the 2000 cross-cutting review implicitly intended to carry out an “output” evaluation and, where appropriate, an “impact” evaluation (HM Treasury, 2000a : 2), although it was not clearly announced in the official documents including the review

report. Hence, the evaluation type in the 2000 cross-cutting review is expected to be appropriate, in that policy evaluation usually aims to measure the direct and short term results of the pertinent policy as well as to investigate both the direct and short term and indirect and long term influences of that policy on the economy and society as a whole.

However, in the 2000 cross-cutting review of science and research, it seems that the impact evaluation may focus on their relatively direct and short term influences, in the light of the fact that this review is in principle to be conducted every two years. This period of two years is not sufficient to achieve fully the original objectives of the policy or of the programmes set up at the beginning of policy formulation.

In consequence, in general, it seems that the overall evaluation paradigm of the 2000 cross-cutting review of science and research was rationally and appropriately established, even though a few adjustments still need to be made, in particular, in relation to the evaluation purposes.

#### ***6.2.1.2. Evaluation Resources***

Since the quantity and quality of evaluation resources can seriously affect the quality of evaluation performance and evaluation findings, it is quite important to examine them in a meta-evaluation. As already discussed, in this meta-evaluation, the evaluation resources are of the following three types : (1) an evaluation manpower, (2) an evaluation unit and (3) evaluation information.

(1) Evaluation Manpower : Two groups of manpower were involved in this review : one was the staff of the Enterprise Team in HM Treasury who did it as their permanent work, while the other type was the external advisers and experts from the DTI, the DfES, the universities and private sectors who did it as temporary work for a period of some months. Of these, one member of staff in the Enterprise Team mainly prepared the evaluation plan and implemented the evaluation process, even though he received a great deal of assistance from the external advisers and experts. In addition to the evaluation work, the staff of the Enterprise Team have carried out many other missions (HM Treasury, 2000 : 3). From this, we can see that the evaluation manpower for this review was not sufficient, in that the 2000 cross-cutting review covered diverse policies and programmes of many

government departments including the DTI, the DfES, the DEFRA, the DTLR and the DoH.

(2) Evaluation Unit : The Enterprise Team in HM Treasury, one of the 19 spending teams overseeing the work of individual departments, played a role as the evaluation unit in the 2000 cross-cutting review. This team was designated to the evaluation unit to take part in many missions including (1) ensuring that public funding concentrates on excellence and rewards exploitation, (2) encouraging universities and public sector research establishments to increase their ability to manage and exploit intellectual property effectively and (3) encouraging more innovation in business (Cabinet Office, 2002a : 2 ; HM Treasury, 2001a : 3). Therefore, it seems that the assignment of the enterprise team as the evaluation unit was effective in the conduct of the 2000 cross-cutting review of science and research, because this team had actual responsibility for the overall management of science and innovation policy at the national level.

(3) Evaluation Information : For the 2000 cross-cutting review, information was mainly drawn from two kinds of sources : one was existing research conducted by external experts mainly belonging to the universities and the other was extensive consultation with senior university academics and the staff of both the Higher Education Funding Councils (HEFCs) and the Wellcome Trust (HM Treasury, 2000a : 3). As a result, much more use was made of qualitative, sometimes slightly subjective, information than of quantitative information. This situation is more or less inevitable in that the policies and programmes normally deal with many complicated matters.

Consequently, it appears that slightly less evaluation resources were invested in the 2000 cross-cutting review. In particular, the evaluation manpower was not sufficient, bearing in mind that the evaluation findings were expected to be utilised as one of the principal references for the 2000 SR and the 2000 Science and Innovation White Paper.

#### **6.2.1.3. Evaluation Performance**

Evaluation performance refers to the substantial elements during the actual evaluation process. In this component, (1) the evaluator, (2) evaluation methods and (3) evaluation criteria and indicators need to be examined. Meanwhile, since the components of the



evaluation performance have been recognised as the core factor of meta-evaluation, there is much research on this subject (Larson and Berliner, 1983 ; Kim, 1993).

(1) Evaluator : This review was actually conducted by the “steering group” under the leadership of the Minister for Science and Innovation in the DTI. The steering group consisted of five members of staff from each of HM Treasury, the DTI and the DfES as well as some external advisers from the universities and the private sector (HM Treasury, 2000a : 3 ; HM Treasury, 2001a : 3). Thus, it is natural that external evaluators played more important roles than internal ones did, since this review handled the S&T and R&D affairs connected mainly with the DTI and DfES, which were the principal policy-makers in the UK SET base.

(2) Evaluation Method : The steering group reviewed the available research literature on the connections between science and economic growth. It also conducted an extensive consultation in order to seek the views of many senior academics in universities, the Committee of Vice Chancellors and Principals (CVCP), the Higher Education Funding Council for England (HEFCE) and the Wellcome Trust (HM Treasury, 2000a : 3). Hence, in this review, the main evaluation method adopted was qualitative. In particular, in conducting the literature review, the steering group relied mainly on second hand materials drawn from existing research, rather than first hand materials.

In general, the evaluation method is considered as one of the important components of evaluation practice, because it will affect the quality of the final evaluation findings and the extent of the realisation of the evaluation purposes (Gibbons and Geroghiou, 1987 ; Miller, 1992 : 3-7 ; Kostoff, 1994 ; Geisler, 1994 : 190-191 ; Yi and Kang, 2000b : 165). For this reason, the evaluation methods for the 2000 cross-cutting review were considered to adopt relatively simple approaches, bearing in mind the extent of the evaluation purposes was multiple and the scope of the evaluation objects was broad.

(3) Evaluation Indicators : In the 2000 cross-cutting review, the steering group adopted four types of evaluation indicators : producing people with knowledge, generating and disseminating new knowledge, the degree of concentration on RAE funding on the best facilities and commercialisation and collaboration with industry (HM Treasury, 2000a : 7, 13, 22). Therefore, both qualitative and quantitative indicators were simultaneously

adopted, even though the usage ratio between them and their weights were different.

In sum, it is clear that, in conducting the 2000 cross-cutting review of science and research, the greater adoption of qualitative approaches reflected the evaluation purposes and the evaluation types, which were defined in a more qualitative way.

#### **6.2.1.4. Evaluation Utilisation**

The evaluation utilisation relates to the feedback process whereby the findings of a primary evaluation will affect the next stages of the policy process. According to existing research, the utilisation of evaluation findings can be classified into two types : instrumental (short-term and direct) utilisation (Alkin, Daillak and White, 1979 : 225-226) and conceptual (long-term and indirect) utilisation (Rich, 1977 : 200 ; Luukkonen and Ståhle, 1990 : 361-363 ; Rossi and Freeman, 1993 : 443 ; Vedung, 1995 : 375-376). Accordingly, firstly, the production and dissemination of the evaluation report will now be discussed and then the types of evaluation utilisation will be examined.

(1) Evaluation Report : The final report of this review was published and can be found on the website of HM Treasury. Its core contents were also summarised in the Spending Review (SR) 2000 White Paper (HM Treasury, 2000a). From this, it appears that the final report of the 2000 cross-cutting review was prepared and disseminated appropriately.<sup>1</sup> However, it seems that there is some possibility for improvements in some elements of the evaluation report. In particular, it is expected that the evaluators, the evaluation methods and evaluation criteria adopted need to be discussed much more, since these components are considered to influence the quality and credibility of the final evaluation findings.

(2) Utilisation Type : This review made a number of recommendations, most of which were reflected in the SR 2000 settlement and the Science and Innovation White Paper 2000 (HM Treasury, 2000a : 3-6 ; HM Treasury, 2001a : 3 ; DTI, 2000b). Therefore, it is clear that the findings of this review were fully adopted as both instrumental and

---

<sup>1</sup> Despite this appropriateness of the review report, there has been a criticism in relation to the delay of its actual publication. In the case of the 2002 cross-cutting review, the final report was published on 22 October 2002, although the *2002 Spending Review White Paper* was announced on 15 July 2002.

conceptual policy means. While inputs to the SR 2000 settlement were short term policy means for the R&D programme and science and research policy, those to the Science and Innovation White Paper 2000 were seen as relatively long-term policy developments in the area of SET policy.

Consequently, it appears that the utilisation of the findings of the 2000 cross-cutting review was rational and appropriate, in that the final report was fully open to the public and the evaluation findings were properly utilised from both the instrumental perspective and the conceptual one.

#### **6.2.1.5. *Synthesis***

In general, the cross-cutting review of science and research is conducted to examine the overall effectiveness and performance of SET activities supported by many organisations in the public sector. Therefore, this review can be considered as typical of policy evaluation in the area of S&T and R&D.

In the light of this, in the meta-evaluation of the 2000 cross-cutting review, it seems that there are both strengths and the things that require improvement in any future exercise. Regarding strengths, the evaluation paradigm was set up relatively satisfactory and the evaluation findings were rationally utilised. Regarding required improvements, the input of evaluation resources, in particular, the evaluation manpower and information, was not sufficient and the evaluation performance needs to be improved, especially since some fairly simple evaluation methods, like the review of secondary sources of information, were adopted.

Therefore, it seems that more diverse evaluation methods which have already been utilised and verified in other evaluation cases need to be adopted (Geisler, 1994 ; Georghiou, 1998 ; Georghiou and Roessner, 2000) as a part of the efforts to improve the quality of the evaluation exercise and to more effectively realise the evaluation purposes set up at the beginning stage.

### ***6.2.2. Meta-evaluation of the Quinquennial Review of the Six Grant Awarding Research Councils in 2001***

In 2001, both the DTI and the OST conducted the first quinquennial review, consisting of two stages, of the six grant awarding research councils, namely the ESRC, the BBSRC, the MRC, the EPSRC, the NERC and the PPARC, to judge the roles and organisational status of the research councils and to examine the efficiency and effectiveness of their operations. In what follows, the rationality of the evaluation system and the effectiveness of the actual practice of the 2001 quinquennial review of six research councils were investigated.

#### ***6.2.2.1. Evaluation Paradigm***

(1) Evaluation Purpose : The 2001 quinquennial review was conducted in two stages whose purposes were slightly different. While the purpose of the first stage was to examine “the role and organisational type” of the research councils, by reference to their Royal Charters, missions and evidence of past practice, the purpose of the second stage was to judge “the efficiency and effectiveness” of the councils’ operations and look for further opportunities for improving performance (DTI/OST, 2001c : 11). As a result, in the 2001 quinquennial review, the government intended to achieve two clear evaluation purposes : “the accountability” of the research councils and the “improvement” of their management. In addition to this, it appears that the “knowledge enhancement” in relation to R&D activities in the public sector was also one of the evaluation purposes, even though it was not expressed in an explicit and direct way in either the review plan or the evaluation report. Therefore, it is clear that the evaluation purposes of the 2001 quinquennial review properly included the general purposes of institution evaluation.

(2) Evaluation Object : The evaluation object of the quinquennial review can be inferred from the evaluation purposes discussed above. It is apparent that while the scope of the evaluation object included a retrospective examination of “past performance” as well as a forward-looking review of the “future management plan”, the level of the evaluation object was confined to the operational aspect, rather than extending to strategies and high-level policies (DTI/OST, 2001c : 12). Therefore, the scope and the level of the

evaluation object were appropriately established, even though more emphasis was placed on the systems and processes used by the research councils and the OST than on the content of their policies and strategies, because of the restriction of the timescale.

(3) Evaluation Type : In the 2001 quinquennial review, both an “output” evaluation and “assessment” of future management were given emphasis. It is clear that the adoption of two types of the evaluation was needed to realise the evaluation purposes mentioned above.

However, it is not reasonable that the conduct of the quinquennial review of research councils should be confined only to the matters relating to past performance and future plans. This is because, in general, the UK research councils have a great deal of influence on the science and research community, including the universities and PSREs. Therefore, it is reasonable for an impact evaluation to be conducted in order to measure and assess the effect and influence of the overall activities of research councils and their performance on the science community as well as on the society as a whole, from the mid and long-term perspectives (Hills and Dale, 1995 : 43).<sup>2</sup>

Consequently, in general, it appears that the evaluation paradigm of the 2001 quinquennial review was proposed relatively rationally and appropriately, even though there is a need to diversify the evaluation type, by for example officially introducing an impact evaluation exercise.<sup>3</sup>

#### **6.2.2.2. Evaluation Resources**

(1) Evaluation Manpower : Three groups of evaluation manpower with different missions

---

<sup>2</sup> Regarding this subject, there may be an opinion that, in the 2001 quinquennial review of six research councils, an impact evaluation was conducted, even though it was not clearly announced in the related formal documents, because the final evaluation findings were also utilised at the conceptual level. However, whether the impact evaluation was conducted officially or not and which type of evaluation utilisation was actually adopted is still a matter of discussion.

<sup>3</sup> The introduction of an impact evaluation in the quinquennial review of the research councils was one of the key recommendations of the final report. According to this, the OST and the research councils should devise a new performance measurement system to reflect how the research councils contribute the science base and which organisational and management structures best deliver this. For this purpose, strategic corporate indicators, one of the four types of indicators for the new measurement system, should measure impact in relation to the achievement of key objectives including the quality of life, wealth creation, the effectiveness of public policy and practice, and science communication (DTI/OST, 2001c : 72, 84)

were involved in the 2001 quinquennial review. This review was overseen by a “steering group” that consisted of 17 members including six council Chief Executives. The actual review was conducted by an independent “review team” comprising six members who came from related government departments, research councils, research charities and business. In addition, in the second stage review, four “working groups” were organised to examine the following four themes : (1) missions, structure and governance, (2) relations between the research councils and their stakeholders, (3) the methods of priority setting and decision-making and (4) the councils’ management and internal processes. The working groups had a wide membership drawn from academia, business, the research charities, government, the research councils and elsewhere and included all 85 members (DTI/OST, 2001c : 11, 13). As a result, a total of 97 members were fully or partially involved in the evaluation. Therefore, it can be said that the government made an effort to engage as wide a range of stakeholders as possible in this review.

(2) Evaluation Unit : The “mini Directorate” which was temporary established under the Science and Engineering Base Group (SEBG) in the OST acted as the evaluation unit, even though many organisations and external experts were involved in this review (OST, 2000a ; OST, 2002). Since the main role of the SEBG is to ensure that the research councils and other publicly funded bodies work efficiently and effectively, it is no wonder that this group was designated to the evaluation unit of the quinquennial review of six research councils.

(3) Evaluation Information : In this review, three kinds of information source were used : (1) extensive consultation with key stakeholders, (2) the related literature and existing studies and (3) an in-depth discussion with staff from the research councils (DTI/OST, 2001c : 12-14). So, much more qualitative than quantitative information was utilised as the main source for the review.

Consequently, it is clear that the overall inputs of evaluation resources to the 2001 quinquennial review were sufficient. Among these resources, in particular, it seems that the involvement of evaluation manpower with a diverse background was appropriate and the types of the evaluation information used were also proper in the light of the evaluation purposes and the evaluation types.

### **6.2.2.3. Evaluation Performance**

(1) Evaluator : Of the three groups of evaluation manpower, a few members of staff in the DTI and the OST were actual responsible for the entire process of the evaluation work. However, in the second stage of the review when the effectiveness and efficiency of the research councils' management were discussed, external advisers from both academia and other related government departments outnumbered internal members of staff from the DTI, the DfES and the research councils (DTI/OST, 2001c : 13). So, it appears that the external evaluators were largely responsible for conducting this review, even though the internal evaluators provided them with much useful information and actual assistance. It was inevitable that the conduct of the 2001 quinquennial review was led by external evaluators, because this review related directly to the overall matters of the six research councils.

(2) Evaluation Method : Three kinds of evaluation method were adopted : (1) public consultation with key stakeholders, (2) review of the related literature and studies and (3) in depth discussion with the members of staff of the research councils. Firstly, in each stage of this review, both the review team and the working group undertook a substantial consultation exercise involving key stakeholders (DTI/OST, 2001b : 3). There was also a written public consultation for stage two of the review (DTI/OST, 2001c : 14, 103). Secondly, a number of related reviews were carried out or had been commissioned in parallel with this review (DTI/OST, 2001c : 11-12). Also, the review team conducted in-depth interviews and regular discussions with the staff concerned in the research councils (DTI/OST, 2001c : 13-14). Therefore, it appears that qualitative analytical approaches were adopted as the main evaluation method.

(3) Evaluation Criteria and Indicators : In this review, the evaluation criteria were not clearly described. However, they can be inferred from the evaluation purpose. We may consider the three following elements as main evaluation criteria : the “appropriateness of the research councils' organisational status”, the “efficiency of the research councils' management” and the “effectiveness of the research councils' management” (DTI/OST, 2001b ; DTI/OST, 2001c). Therefore, it is apparent that more qualitative criteria were adopted than quantitative ones, although both types were utilised in the actual evaluation

process. Meanwhile, there was also no clear explanation of the evaluation indicators in the related official documents. As a result, it seems that the greater emphasis on the qualitative criteria has something to do with the evaluation paradigm, in particular, the evaluation purposes and the evaluation types being defined in a relatively qualitative way.

Therefore, we may conclude that the components of the evaluation performance were rationally adopted in order to accomplish the initial evaluation purposes set up in advance. In particular, it is clear that the adoption of the diverse evaluation methods to represent the opinions of many stakeholders was excellent. This is because the final findings of the 2001 quinquennial review will affect research communities, including the research councils, their research institutes and universities, in various ways and have different influences on them.

#### **6.2.2.4. *Evaluation Utilisation***

(1) Evaluation Report : The reports of stage one and two of the review were published (DTI/OST, 2001b ; DTI/OST, 2001c) and are also available on the OST website. In particular, the report of stage two of the review, regarded as a final report of the 2001 quinquennial review, dealt with some issues regarding the future development of research councils and made many recommendations concerning these (DTI/OST, 2001c). Therefore, it is clear that the evaluation report was regarded as a good, one from the point of view of its production, its detailed contents and dissemination to the related people and institutions.

(2) Utilisation Type : The principal recommendation of stage one was that the research councils should continue to be the NDPBs (DTI/OST, 2002b : 19). The major conclusions and recommendations of stage two were one of the most important inputs to the 2002 cross-cutting review of science and research, a part of the 2002 spending review (DTI/OST, 2002c : 5-10). So, the related government departments, including HM Treasury, the DTI, the DfES and the OST, adopted these recommendations as a future policy means (HM Treasury et al., 2002). Therefore, it seems that the evaluation findings were fully utilised to improve the research councils' management systems as well as, in the broader perspective, to re-arrange their roles and missions in line with UK science and technology policy. As a result, it is clear that the utilisation of the final evaluation findings



was very appropriate from the perspective of both instrumental and conceptual utilisation.

#### **6.2.2.5. *Synthesis***

In general, in the UK, the quinquennial review of the agencies and the NDPBs is conducted in order to evaluate their roles and organisational status as well as to examine the efficiency and effectiveness of their current management system. So, it appears that this review normally includes both the past performance of the pertinent organisations and the future plans and strategies for development.

For this reason, it is essential that the evaluation paradigm, in particular the evaluation purposes, of the 2001 quinquennial review of the six grant awarding research councils was established in broader term. It is apparent that this affected decisions that were made about the extent of evaluation resources invested and which kind of the evaluation performance was adopted. Therefore, it seems that the rate of investment of the evaluation resources was reasonable and the contents of the evaluation performance adopted were suitable for the realisation of the evaluation paradigm. Meanwhile, the final evaluation findings were utilised fully to reflect the evaluation purposes of the 2001 quinquennial review.

Consequently, it is believed that the evaluation system of the 2001 quinquennial review of the six grant awarding research councils was set up rationally and that its exercise was also performed effectively. However, it seems that there may be room for improvement, in particular, in relation to the evaluation types, in the sense that an impact evaluation was neither actually conducted nor clearly announced in the official documents and related guidelines.

#### **6.2.3. *Meta-evaluation of the Institute Assessment Exercise (IAE) by the BBSRC in 2001***

The BBSRC supports research in its sponsored research institutes. It provides funding for research institutes through Core Strategic Grants (CSG) on the basis of an extensive

four-yearly review of past performance and future plans, namely the “Institute Assessment Exercise (IAE)”. The 2001 IAE was the second exercise of this type, following the first IAE in 1996-1997.

#### **6.2.3.1. Evaluation Paradigm**

(1) Evaluation Purpose : It was officially announced that the BBSRC had conducted the 2001 IAE in order to allocate the CSG to the eight institutes it directly sponsored and the Horticulture Research International (HRI) under the responsibility of the DEFRA for the four financial years from 2002-2003 to 2005-2006 (BBSRC, 1999a : 1). From this statement, it is apparent that the emphasis of the evaluation purposes of the IAE 2001 was placed firstly on the “accountability” of individual institutes and that “improvements” of the expenditures of the research fund were considered as an additional evaluation purpose.

However, there was no clear identification of the knowledge enhancement in relation to R&D activities in public sector. In actual fact, in the 2001 IAE practice, individual institutes have played a principal role in the areas of biotechnology and the biological sciences, by carrying out long-term and strategic R&D programmes in these areas. Therefore, it is expected that, through conducting the IAE regularly, the existing theories should be confirmed and new ones developed in these science and research areas.

(2) Evaluation Object : The 2001 IAE comprised the four following assessments : assessments by expert panels on the quality of science (QS), an assessment by a panel on knowledge transfer (KT), an assessment by the BBSRC office of Competitive Funding (CF) and an assessment of the institute as a whole by a visiting group (VG) (BBSRC, 1999a : 1). On the basis of these, we may infer that there were four evaluation objects : (1) the quality of science, (2) knowledge transfer, (3) competitive funding and (4) the overall activities of the institute. While some objects such as the evaluation of knowledge transfer and the visiting group assessment covered both the “past performance” and the “future research plans” of individual institutes, other objects such as the assessment of the quality of science and the competitive funding included only past performance. Thus, it is apparent that the selection of the evaluation objects was appropriate, in that they covered the major results of the R&D activities as well as matters relating to the management of

individual research institutes.

(3) Evaluation Type : The evaluation type can be inferred from the evaluation purposes and evaluation objects discussed above (BBSRC, 1999a : 1), although there was not an explicit announcement concerning these. Consequently, in the 2001 IAE, an “output” evaluation and the “assessment” of the research plan were conducted in each institute simultaneously. It is believed that, as a result, more emphasis was placed on the output evaluation of the four evaluation objects than an assessment of the knowledge transfer (KT) and the visiting group (VG) assessment.

As in the case of the meta-evaluation of the evaluation purpose, it is not proper that the 2001 IAE was conducted only as an output evaluation and the assessment of future plan. Therefore, it is necessary for an impact evaluation to be introduced, in order to judge the influences of the each institute’s R&D activities on the related science community and to assess the impact of the research programmes carried out by the institutes on development in the areas of biotechnology and the biological sciences. In addition, the need to carry out an impact evaluation will be higher, bearing in mind that the IAE practice is normally conducted every four years.

In consequence, it seems that the evaluation paradigm of the 2001 IAE needs to be properly adjusted and extended. In particular, it is rational that the evaluation purposes should be expanded to cover the enhancement of knowledge and that an impact evaluation should also be introduced in line with the extension of the evaluation purposes.

#### **6.2.3.2. *Evaluation Resources***

(1) Evaluation Manpower : Two groups of evaluation manpower were involved in the 2001 IAE : the external expert panels, including visiting groups, and the officers of the BBSRC. The six panels, each of 12 members from academia, business and government, carried out the evaluation of the quality of science. Another single panel, consisting of 12 members drawn mainly from industry, carried out the evaluation of knowledge transfer (BBSRC, 1999a : 3). The nine visiting groups per individual institutes, each consisting of 12 members from UK universities, business, government and where appropriate from overseas, considered the totality of the institute’s research for the previous four years and

the institute's forward plans for the following four years (BBSRC, 1999a ; BBSRC, 2002a). Meanwhile, members of staff in the Finance Group, one of the four departmental groups in the BBSRC, examined the external income from open competition from April 1996 to March 2001 (BBSRC, 1999a). As a result, a total of around 190 external experts and internal staff with high expertise were involved in the 2001 IAE.

(2) Evaluation Unit : The evaluation unit was the "Assessment and Policy Unit" (APU), consisting of six members of staff as at January 2002, which operated within the "Business Innovation and International Group" (BBSRC, 2002a). Unlike in other cases, this unit carried out the actual evaluation of some evaluation objects as an internal evaluator, alongside its function as the secretariat in the 2001 IAE (BBSRC, 1999a). Thus, the APU coordinated the overall activities of the 2001 IAE, with assistance from the Science and Technology Group in the case of the QS, the Finance Group in the case of the CF and the Business and Innovation Unit in the case of the KT.

(3) Evaluation Information : The evaluation information came mainly from the performance report submitted by individual institutes. It included both qualitative and quantitative information, depending on the evaluation objects (BBSRC, 1999a). From this, it is clear that both quantitative and qualitative information was adopted in the 2001 IAE, even though the detailed information requirements differed according to the evaluation object. One of the features of the evaluation information is that much more quantitative information was required than in other UK evaluation exercises. It is assumed that the greater utilisation of quantitative information had something to do with the evaluation methods and the evaluation indicators adopted during the evaluation performance.

As a result, it can be said that the overall inputs of the evaluation resources were sufficient and suitable for the realisation of the evaluation purposes. Also, the quality of the evaluation resources, in particular, the evaluation manpower and evaluation information were relatively higher than in other UK exercises.

### **6.2.3.3. *Evaluation Performance***

(1) Evaluator : In the 2001 IAE, two types of evaluator were engaged : the BBSRC

officers such as the staff of the APU and of the Finance Group, and the external experts involved in both expert panels and visiting groups (BBSRC, 1999a). However, the external evaluators played more important roles in this evaluation, in a sense that they evaluated the QC, the KT and the VG, which were regarded as the main evaluation objects.

(2) Evaluation Method : Qualitative methods such as site visits and report reviews as well as quantitative methods including bibliometrics were adopted. However, quantitative evaluation methods were utilised more than in other evaluation exercises, including the 2000 cross-cutting review and the 2001 quinquennial review. In conducting the retrospective assessment of the QS and the KT, quantitative methods including numbers of publications and awards, and output and performance indicators (OPIs) were mainly used. In particular, in evaluating the quality of the published research articles, a bibliometric analysis was used in order to measure scientific and technical excellence (BBSRC, 1999a). Meanwhile, in assessing the future prospect for the KT and the institute's performance as a whole, a qualitative evaluation method was adopted. Thus, when the visiting groups evaluated the overall performance of the institutes, they spent three days at each institute (BBSRC, 2002d).

(3) Evaluation Criteria and Indicators : In the evaluation of some objects such as the QS, the CF and the retrospective KT, a quantitative measurement unit was used as the major evaluation indicator. On the other hand, the qualitative criteria had to be utilised, in the prospective assessment of KT and the VG (BBSRC, 1999a). Therefore, in the 2001 IAE, it appears that quantitative evaluation indicators were adopted as much as possible, where available, depending on the evaluation objects.

Consequently, it appears that the detailed components of the evaluation performance were rationally adopted and properly implemented. In particular, it is thought that the much greater adoption of the quantitative evaluation methods and indicators reflected the attributes of the evaluation objects and evaluation information, which were defined in a more quantitative way.

#### **6.2.3.4. Evaluation Utilisation**

(1) Evaluation Report : The final report of the 2001 IAE was not available to the public in either an electronic form or hard copy. The summary of the outcome of this evaluation was merely posted on the BBSRC website (BBSRC, 2002d). Consequently, this report needs to be published and posted on the website as in other evaluation cases, bearing in mind that it is expected to function as an important reference for the BBSRC-supported institutes as well as for the research communities concerned in the future.

(2) Utilisation Type : The BBSRC Council used the evaluation findings of the 2001 IAE to set funding plans for the period 2002-2006 and endorsed the key recommendations. Firstly, on the basis of the results of the IAE 2001, the BBSRC Council agreed changes to the institute's CSG and adjusted the total values of the responsive mode awards for individual institutes (BBSRC, 2002b ; BBSRC, 2002d ; BBSRC, 2002e). Secondly, the evaluation findings were also devoted to the developments in the management system and to the research strategic plan of the BBSRC institutes and of the HRI. Thirdly, they may influence the Councils' ten-year scientific opportunities road map and the next five-year strategic plan. From this, it appears that the outcomes of the 2001 IAE fed properly into the BBSRC's strategic planning and research arrangements (BBSRC, 2002e). As a result, the findings of the 2001 IAE were fully accepted as a policy means in relation to budget allocation (instrumental utilisation) as well as to the long-term developments in the management system and the research plan for both the related institutes and the BBSRC itself in the future (conceptual utilisation).

In consequence, it is clear that the type of utilisation of the final findings of the 2001 IAE was very rational and appropriate. However, it is no doubt that there is a great need for improvement concerning the dissemination of the evaluation report. Thus, the evaluation report should be open to the related research community as well as to the public.

#### **6.2.3.5. Synthesis**

In case of the 2001 IAE practice, in general, it is clear that sufficient evaluation resources were invested and that the evaluation performance was also rationally adopted and effectively implemented. However, in order to enhance the quality of the evaluation

exercise, there remain some improvements that need to be made, in relation to diversification of the evaluation types and the dissemination of the final evaluation report. In particular, there is a serious need for these improvements, in the light of the fact that the IAE practice normally deals with many things concerning the research programmes as well as R&D management matters for the period of at least four years in the areas of biotechnology and biological sciences.

#### ***6.2.4. Meta-evaluation of the Evaluation of Funding for UK Civil Space Activity by the DTI in 2001***

Between late 1999 and early 2001, the five funding partners, namely the DTI, the DETR (now DfT), the Met. Office, the NERC and the PPARC, conducted a cross-departmental evaluation of UK civil space activities. This was the first comprehensive evaluation of the impact and achievements of public support for UK civil space activities. Initially, this evaluation was launched by the Evaluation and Policy Improvement Committee (EPIC) of the DTI. However, the evaluation was extended to encompass the full range of UK public funding on civil space activities, in order fully to consider in equal measure the industrial, scientific and public policy benefits that had been realised. This evaluation was carried out on the basis of the comprehensive self-assessments performed by the five funding partners, using the common evaluation protocol prepared by the evaluation team. Meanwhile, the evaluation methodology adopted here was novel in as much as it combined self-assessment with a separate validation procedure involving R&D evaluation professionals and international space experts.

The meta-evaluation was conducted only on the DTI's evaluation exercise, even though the five funding partners were involved in this evaluation. This is because the DTI was the initial organiser of this evaluation and also undertook a more important task during the evaluation process.

##### ***6.2.4.1. Evaluation Paradigm***

(1) Evaluation Purpose : The DTI evaluation was intended to examine the rationale for and the impacts and achievements of public funding of all UK civil space activities.

Additionally, policy makers needed to consider how future policy on space should be delivered and managed in order to maximise the potential gain from this activity for the UK and international space supply chains (DTI/AU, 2001 : 138, 139). Therefore, it is clear that ensuring the “accountability” of past expenditure was the most important evaluation purpose. In parallel, it appears that “programme improvement” in the space sector was also conceived as a main evaluation purpose.

Consequently, it is evident that the main purposes of the DTI evaluation were rationally proposed. However, this evaluation needs to include knowledge enhancement as a further evaluation purpose. Since this work was a complex and wide-ranging evaluation covering a long period of time, namely some 25 years of ESA membership, and both international and domestic activities (DTI, 2001c : ii), it is expected that related policy makers and researchers will obtain some useful knowledge and information for future theoretical developments in the space sector.

(2) Evaluation Object : As already pointed out, this evaluation covered complex and wide-ranging affairs implemented by many government departments and included both international and domestic activities (DTI/AU, 2001 : 137). From this, it is clear that the level of the evaluation object normally covered the various “programme-level works” including pure science and R&D programmes, the application of the research outputs and international collaboration in the space sector. Additionally, it sometimes included “policy-level work”, when one or more government departments supported a specific activity. Meanwhile, the scope of the evaluation object covered “home programmes” as well as “international programmes”, including collaborative R&D programmes between ESA members. Therefore, it appears that the evaluation objects were selected in a comprehensive way and that their level and scope could be also controlled.

(3) Evaluation Type : Officially, the DTI evaluation was undertaken to judge the impacts and achievements derived from UK public funding of civil space activities as well as to provide the policy makers concerned with the programme improvements necessary for the future (DTI/AU, 2001 : 138, 139). From this, it is clear that the DTI exercise was carried out both as an “output” and “impact” evaluation and also as a prospective “assessment” of the programme improvement. Meanwhile, regarding outputs and impacts, the overall cross-departmental evaluation examined, in equal measure, the



industrial, scientific and public policy benefits, whereas the DTI evaluation focused on the industrial and commercial aspects of public funding (Technopolis Group, 2001 : 11 ; DTI/AU, 2001 : 138). As a result, it is clear that the evaluation type represented by the DTI exercise was adopted properly in accordance with the initial evaluation purposes, since these covered all evaluation types that have been considered as typical of programme evaluation.

In sum, it appears that the overall evaluation paradigm of the DTI exercise was set up more rationally than in the case of other evaluations in the UK. In particular, both the selection of the evaluation objects and the adoption of the evaluation types properly reflected the initial evaluation purposes. However, as in the other cases, it is necessary for the knowledge enhancement to be added as an additional evaluation purpose, even though it will be described a somewhat broadly in related formal documents.

#### ***6.2.4.2. Evaluation Resources***

(1) Evaluation Manpower : Two groups of manpower were involved in the DTI's evaluation : the first was the external evaluation team and the second was the internal one. The external evaluation team was awarded the contract for this work by open competition and comprised five specialist policy and evaluation consultants from the Technopolis Group and three international space experts. The main roles of the evaluation team were to provide advice and guidance on aspects of evaluation methodology to the internal evaluators including the Assessment Unit (AU) in the DTI as well as to draw conclusions and make recommendations for appropriate improvements (DTI, 2001c : 7 ; Technopolis Group, 2001 : 127). Meanwhile, the three evaluators in the DTI conducted the actual evaluation of the DTI's funding programmes for space under the guidance of the evaluation team and, sometimes, in collaboration with them (DTI/AU, 2001).

As a result, a total of eight people were involved in the DTI evaluation. Therefore, it is clear that the input of evaluation manpower was appropriate and high, in that those who conducted the evaluation belonged either to an independent directorate of the DTI or the external consultation company both of which have a long and a prominent experience in the area of R&D evaluation.

(2) Evaluation Unit : This evaluation was organised and carried out by the Assessment Unit, one of two units under the Technology, Economics, Statistics and Evaluation (TESE) Directorate in the DTI. As at February 2002, the AU comprised five professional evaluators most of whom had a scientific or technological background. In carrying out evaluation, the AU drew on the expertise of four professional economists in the TESE (Barber, 1999a : 6 ; DTI, 2002a). From this, it appears that the AU consisted of members of staff with the high-quality expertise necessary for the evaluation of science and technology programmes.

(3) Evaluation Information : The evaluation information came mainly from the following five sources : (1) a newly commissioned study for this evaluation exercise, (2) meetings with the British National Space Centre (BNSC) programme managers, other experts and peers concerned, (3) extensive structured interviews with direct funding beneficiaries, (4) a review of archived files and (5) a review of the cases of other ESA member (DTI/AU, 2001 : 140). From this, it is apparent that more diverse and plentiful information sources were utilised to sort out the key evaluation questions. Also, it seems that qualitative information was adopted as the main source of information for evaluation, although some quantitative information was utilised in the actual process of evaluation where appropriate.

In consequence, it is thought that the investment of the evaluation resources was determined at a higher level than in other cases. In particular, the input of evaluation manpower and evaluation information was sufficient and adequate in both quality and quantity. Another feature of the evaluation resources invested is that the role of the evaluation unit was performed by the AU, which is a standing, independent directorate in the DTI, whose principal mission is to conduct the overall assessment, including the *ex post* evaluation of the S&T programmes supported by the DTI.

#### **6.2.4.3. Evaluation Performance**

(1) Evaluator : The three evaluators under the AU took completely responsibility for the evaluation of the space programmes funded by the DTI, even though the two external consultants, Technopolis and Scientific Generics, were fully involved in the five funding partners' evaluation through preparing the overarching report and conducting a mapping

study of the UK space, respectively (DTI/AU, 2001). Therefore, it is apparent that, in fact, the internal evaluator conducted the DTI evaluation. In contrast to other evaluation exercises normally conducted by external evaluators, it is a feature of the DTI evaluation that it was undertaken by internal evaluators possessing high expertise in the area of R&D evaluation.

(2) Evaluation Method : The AU adopted diverse evaluation methods as follows : (1) a field survey in the manner of mapping exercise, (2) a review of the literature including a newly commissioned study and archived files, (3) an in-depth interview and (4) a comparative study between the UK and other ESA countries (DTI/AU, 2001 : 214-216). Therefore, it seems that typical types of qualitative method were adopted as the key evaluation methods.

(3) Evaluation Criteria : The main criteria adopted in the DTI evaluation were as follows : (1) the “rationale” of the government intervention in civil space activity and its current “relevance”, (2) the “additionality” of the achievements of public funding and the existence of “alternatives”, (3) the principal “beneficiaries” of public funding, (4) the “appropriateness” of public funding among the supporting programmes, such as those relating to communications, earth observation, science and so on and (5) the “balance” between international programmes and national programmes (DTI/AU, 2001 ; 139). From this, we may see that rather qualitative evaluation criteria were adopted to judge the benefits and impacts of public funding of space activities.

In general, it is thought that the space activities have many characteristics of a public good. So, most benefits and impacts derived from them are difficult to trace, quantify or value. For this reason, it is thought that rather qualitative evaluation methods and criteria were adopted in the DTI evaluation.

#### **6.2.4.4. *Evaluation Utilisation***

(1) Evaluation Report : Unlike in previous DTI evaluation exercises, in which only abstract of the final report was publicly available, the overarching report of the funding partners’ evaluation was fully published on the TESE and BNSC website. It has also been available free on CD-ROM and in hard copy comprising six chapters (DTI, 2001c ; DTI,

2002a). The AU has also published its own separate evaluation report made up of 10 sections and three annexes (DTI/AU, 2001). Therefore, it is believed that the evaluation report was excellently produced and broadly disseminated to the related people and institutions.

(2) Utilisation Type : In the overarching report, Technopolis, an evaluation team, synthesised the findings of the separate evaluations by the funding partners and made various recommendations. All of these issues, except one, were accepted by the government via the BNSC Resources Board (BNSC, 2002 : 1-6). Meanwhile, the AU also made many recommendations in its separate evaluation report, which sometimes slightly overlapped those of the evaluation team. Certain individual recommendations were fully accepted by the government through the DTI (BNSC, 2002 : 7-8). Subsequently, it has become apparent that most of the findings of the civil space evaluation were accepted by the related government departments. So, it is expected that they will be fully utilised both instrumentally and conceptually for the policies and strategies of the government departments, which intend to enhance the space activities under their responsibility.

As a result, it is believed that the utilisation of the civil space evaluation was excellent. It seems that this resulted from the initial intention of the government to judge the past achievements and various impacts of space activities as well as to know how space policy should be delivered and managed in future in order to maximise the potential benefits for the UK industries concerned both at domestic and international level. In addition, the final findings of the DTI evaluation were adopted as a useful reference for the improvements of the space R&D programmes supported by public funding, as well as of the long-term developments of the space related industries of the UK.

#### **6.2.4.5. *Synthesis***

In general, it is clear that the evaluation of UK civil space activities was regarded as the best exercise among the five evaluation case studies. That is, the evaluation purposes were set up to properly cover the performances and benefits of the UK space activities that had been realised during a period of 25 years, even though they placed great emphasis on both accountability and programme improvement. The evaluation resources and the evaluation performance, which were considered as main instrumental factors

during the evaluation process, were also well invested and rationally implemented in order to fully realise the original evaluation purposes. As a result, it is expected that both the sufficiency of the evaluation resources invested and the appropriateness of the evaluation performance implemented were associated with the good production of evaluation findings. Therefore, it is no wonder that the utilisation of the evaluation findings was very good from the perspective of both instrumental and conceptual utilisation.

#### ***6.2.5. Meta-evaluation of the Review of the Gastrointestinal Response to Food (GRF) Programme by the BBSRC in 2000***

The BBSRC has a policy to evaluate all of its research programmes and to use the findings to inform future funding decisions. In line with this policy, in May 2000, it evaluated the “Gastrointestinal Response to Food (GRF)” programme, which had been started in October 1992. The GRF programme has been one of the research priorities in the Agri-Food Committee, one of the seven standing scientific committees in the BBSRC. As at May 2000, a total of 31 grants were awarded to 26 Principal Investigators (PIs) under the GRF programme. The total funding for these awards was £6,434,763 (BBSRC, 2000 : 13).

##### ***6.2.5.1. Evaluation Paradigm***

(1) Evaluation Purpose : In general, the BBSRC evaluates all research programmes in order to assess (1) the achievement of their research objectives, (2) contributions to the related science and research area, (3) benefits and added value from the programme and (4) the level and quality of the research outputs (BBSRC, 2001b). It is assumed therefore that these general evaluation purposes were the terms of reference for the evaluation of the GRF programme. From this, it appears that the evaluation purpose was to secure “accountability” and “programme improvement”. Therefore, it seems that the purposes of the evaluation of the GRF programme were properly established, in that these were normally considered as the principal purposes in the evaluation of R&D projects.

(2) Evaluation Object : At the stage of implementation, the GRF programme comprised

31 research projects carried out by the 26 principal investigators that were more relevant (BBSRC, 2000 : 1, 13). So, the actual evaluation needed to be conducted on the basis of the individual R&D projects, even though, in principle, it was intended to examine the programme itself as a whole. Therefore, it appears that the evaluation objects were the individual projects of the GRF programme.

(3) Evaluation Type : Since there was no clear announcement about the evaluation type, this could be inferred from the evaluation purposes discussed above (BBSRC, 2001b ; BBSRC, 2000 : 1). Therefore, it is clear that, in principle, this evaluation was conducted as an “output” evaluation and additionally as an “impact” evaluation, where appropriate. So, we may see that the types of evaluation of the GRF programme were those typically used in the evaluation of R&D programmes or projects.

Consequently, it seems that the overall evaluation paradigm of the GRF programme was properly proposed. In particular, as in other evaluation exercises, there was also no explicit announcement concerning knowledge enhancement. However, it is thought that some restriction of the evaluation purposes is quite appropriate, bearing in mind that the evaluation was concentrated on the research outputs and the quality of individual projects rather than of the programme itself.

#### **6.2.5.2. Evaluation Resources**

(1) Evaluation Manpower : Two groups of evaluation manpower were involved in the evaluation : “members of staff in the Assessment and Policy Unit” (APU) and “the review panel”. The main role of the APU staff was to issue a questionnaire to all grantholders, to analyse their responses and to conduct a bibliometric analysis of the research outputs (BBSRC, 2001b). The review panel, which consisted of around 12 members from several members of the Agri-Food Committee plus other independent members, considered the questionnaire responses and bibliometric analysis as well as reviewing the scores of individual projects from the final report and sometimes the progress report (BBSRC, 2001c).

(2) Evaluation Unit : In this evaluation, both the “APU” within the Business Innovation and International Group and the “Agri-Food Branch” within the Science and Technology

Group acted as the evaluation units. The APU has principal responsibility for the evaluation of the research institutes and research programmes. As at January 2002, it comprised six members of staff (BBSRC, 2002a). Meanwhile, the “Agri-Food Branch” played a direct part in the evaluation by organising the review panel and acting as the secretariat to the Agri-Food Committee. As at January 2002, there were eight members of staff in this branch and their main role was to arrange and support the review panel, which had taken into deliberate account the results of both the questionnaire analysis and the bibliometric analysis conducted by the APU (BBSRC, 2002a).

(3) Evaluation Information : Three types of source were utilised to gain information for the evaluation : (1) a “questionnaire” returned from the grantholders, (2) a “bibliometric analysis” of the research outputs and (3) “final report forms”. The questionnaires were analysed to gain various information, including the achievement of programme objectives, the BBSRC’s management system, new contacts and collaborations and further funding applied for and received (BBSRC, 2001b). A bibliometric analysis was conducted in order to obtain the information about both the quality and level of research outputs, with emphasis on papers in peer-reviewed journals (BBSRC, 2000 : 13). A final scientific report form must always be submitted within three months of the completion of the grant period. In some circumstances, interim reports or site visits may be required. Generally, final reports are assessed by two suitable experts (BBSRC, 2001b ; BBSRC, 2001c : 15-16). As a result, in this evaluation, while qualitative and quantitative information were equally utilised, quantitative information was considered to be more important than in other evaluation exercises.

In consequence, it appears that, the resources for the evaluation of the GRF programme were sufficient and rational. Also, there are some noteworthy features of the evaluation resources used. First of all, the internal manpower, namely the staff of the APU, played an important mission during the evaluation process, in that they produced the most essential information for the evaluation, by conducting the questionnaire analysis and a bibliometric analysis of the research outputs. In contrast to other UK evaluation exercises, in this exercise, two internal organisations were designated as the evaluation unit, even though their roles in the actual evaluation process were slightly different. Finally, the evaluation information came from various first hand sources, with much more emphasis on quantitative sources than on qualitative ones. As a result, it is clear that these

characteristics of the evaluation resources functioned as an important vehicle for the good implementation of the evaluation performance.

### **6.2.5.3. *Evaluation Performance***

(1) Evaluator : The staff of the APU conducted a questionnaire analysis and a bibliometric analysis, whereas the review panel considered and reviewed the results of the APU's work (BBSRC, 2001b). So, it is clear that the internal evaluators (the APU staff) played a more important role than the external ones, in that they were directly involved in the questionnaire analysis and bibliometric analysis which were adopted as the main methods of evaluation of the GRF programme.

(2) Evaluation Method : Both the "analysis of the grantholders' view" obtained from the questionnaire and the "bibliometric analysis" of the research outputs were adopted as main evaluation methods. The questionnaire was distributed to the 26 principal investigators with the aim of gaining participants' views on a range of various issues (BBSRC, 2000 : 1). For the bibliometric analysis, calculations were based on research outputs given on the final and interim report forms in addition to the supplementary information extracted from the questionnaires. The research outputs considered included papers in peer-reviewed journals, other significant publications and presentations, further grants obtained, public understanding of SET and potential exploitation of research results. In particular, a greater weight was given to papers published in peer-reviewed journals than to those either in press or merely submitted (BBSRC, 2000c : 13-14).

(3) Evaluation Criteria and Indicators : There were no clear descriptions of the evaluation criteria and indicators. However, it was possible to extract these from many official documents including the results of the questionnaire analysis and bibliometric analysis (BBSRC, 2000), the assessment criteria for the final report forms (BBSRC, 2001c : 15-16) and the appraisal criteria for research projects (BBSRC, 1999b : 25). From these references, it is apparent that the achievement of research objectives and scientific excellence were adopted as essential criteria, while improvement of the quality of life and the contribution to the UK industrial competitiveness were considered as further evaluation criteria. Meanwhile, the main evaluation indicators adopted were the size and nature of further funds, frequency of attendance at workshops and meetings, the total



number of published papers, the number of papers in peer reviewed journals and the ratio of refereed publications to all research papers.

In sum, it is believed that the evaluation performance of the GRF programme was adequately and relevantly implemented. In particular, the adoption of the evaluation methods and criteria, which were described in a quantitative way, was well connected to the evaluation purposes and objects, which were defined in a quantitative way.

#### **6.2.5.4. *Evaluation Utilisation***

(1) Evaluation Report : The evaluation report was not made available to the public and was circulated only to the BBSRC office and related staff. This report contained the general information including the programme structure, the questionnaire sheets, a list of peer-reviewed articles and other research outputs (BBSRC, 2000).

(2) Utilisation Type : In the BBSRC, the findings of the evaluation of research programmes have been generally utilised for decision making about whether they should be continued, reduced or terminated (BBSRC, 2002a). In addition, the evaluation findings of the GRF programme were utilised to identify priority areas for future funding (BBSRC, 2001b).

As a result, it appears that the findings of the evaluation of the GRF programme were utilised properly in accordance with the evaluation purposes initially set out. Despite the proper utilisation of the evaluation findings, the evaluation report needs to be more broadly disseminated, in particular, to the outside. This is because the evaluation report is normally expected to provide the researchers concerned with very useful information for research activities, including research planning, in the future.

#### **6.2.5.5. *Synthesis***

Unlike other evaluation exercises such as the 2000 cross-cutting review, the 2001 quinquennial review of research councils and the 2001 evaluation of funding for civil space activities, which were regarded as a large scale evaluation work, the evaluation of the GRF programme in 2000 was on a relatively smaller scale. So, it is expected that this

basic feature of the evaluation task affected the overall evaluation exercise and its detailed process, in particular, the evaluation paradigm and the evaluation utilisation. In line with these characteristics, the evaluation purpose was much concentrated on the accountability of the research leaders and their expenditure of research funds. As a result, it seems that other evaluation purposes such as programme improvement and knowledge enhancement were either ignored or less emphasised than in other evaluation exercises, including the 2001 quinquennial review of six research councils and the 2001 evaluation of civil space activities. In addition, much more emphasis was placed on the direct and instrumental utilisation of evaluation findings, including the decision about the continuity of the pertinent research project and the identification of priority areas for future funding, than on its long-term and conceptual utilisation, such as identifying the future direction of R&D activities through research planning and the analysis of technological trends (Barber , 2001 : 9.30-9.33).

Even though there was some need to adjust the evaluation paradigm and extend the evaluation utilisation, it is clear that the evaluation resources were sufficient and that the evaluation performance was also adequate. In particular, the roles of the internal manpower, namely the staff of the APU and the Agri-Food Branch, were more important than in other evaluation exercises, in that they were fully involved in the actual evaluation, like the secretariat to the review panels. Also, the much greater adoption of quantitative evaluation methods and evaluation criteria was relevant, because this evaluation aimed to examine and judge the research outputs, including the excellent research papers, which were expected to be more easily measured and quantified.

#### ***6.2.6. Section Conclusion***

From the findings of the meta-evaluation of the five UK evaluation exercises, in principle, it seems that the instrumental and methodological components of the R&D evaluation system, for example, the evaluation resources and the evaluation performance, were sufficient and suitable, whereas more fundamental and directive component such as the evaluation paradigm needed to be improved. Meanwhile, the evaluation utilisation, which is regarded as a feedback process for the next stage of the policy process, was relatively appropriate. In what follows, some characteristics and limitations of the UK R&D evaluation system will be discussed by reference to some sub components of the

meta-evaluation model.

Firstly, in three out of the five evaluation exercises, i.e. excluding both the quinquennial review of research councils and the review of the GRF programme, the evaluation purposes need to be adjusted to include knowledge enhancement, even though it will be expressed in a somewhat an implicit and indirect way. This is because these evaluation exercises normally deal with large-scale policies and programmes that have a great deal of long term influence on SET areas. For this reason, an impact evaluation needs to be added to the four evaluation systems, excluding the review of the GRF programme, as a further evaluation type.

While improvements need to be made to evaluation purposes and types, the definition of the evaluation objects was completely satisfactory. In particular, the evaluation objects of the 2000 cross-cutting review and the evaluation of civil space activities were very properly set up, in that these exercises included multiple and diverse policies and related sub programmes carried out by many government departments and the relevant public organisations.

Secondly, in general, it is believed that the evaluation resources used were sufficient and suitable, even though a few improvements are still needed in some evaluation exercises. First of all, one feature of the evaluation resources can be extracted from the composition of the evaluation manpower. As in general cases of R&D evaluation, two groups of evaluation manpower were involved in the five UK evaluation exercises : the internal staff of the evaluation unit and the external experts consisting of the evaluation committee or review panel. The external manpower normally came from academia, industry, the research institutes and government departments concerned. From this, we can see that the policy of the UK government is that the related government departments should participate in the overall evaluation process of their policies and programmes.

In particular, in the five UK evaluation exercises, an independent and professional organisation was designated as the evaluation unit. Meanwhile, in some evaluation exercises, in addition to acting as the secretariat to the evaluation panels, which was regarded as their basic mission, members of the evaluation unit were directly involved in the actual evaluation of some sub evaluation objects. Therefore, this is regarded as one of

the distinctive characteristics in the UK R&D evaluation system, because the high expertise of the evaluation unit will make a great contribution to the quality of the overall evaluation exercise and, in turn, the degree of the evaluation utilisation.

Despite the general strengths of the evaluation resources, in the case of the 2000 cross-cutting review, there are some possibilities for improvement in the evaluation information. In this review, the principal evaluation information came from existing research, which did not include first hand materials but second hand ones. Therefore, much more effort needs to be made in order to get and use evaluation information from a greater variety of sources.

Thirdly, it is thought that the evaluation performance was effectively implemented in relative. In particular, in some evaluation exercises such as the 2001 IAE, the evaluation of civil space activities and the review of the GRF programme, the internal evaluators played more important roles than the external evaluators did. That is, they were fully involved in the actual evaluation, for example, through preparing the common evaluation protocol for the evaluation of civil space activities and conducting the bibliometric analysis, one of the main evaluation methods, of the 2001 IAE and the evaluation of GRF programme.

However, there is also a need for improvement in the evaluation methods, in particular, in the case of the 2000 cross-cutting review. As already noted, in this review, a rather simple evaluation method, in the form of a review of the small amount of existing research available, was adopted. Therefore, it seems that more diverse evaluation methods, which had already been utilised and verified in other evaluations, need to be adopted, bearing in mind that, in the cross-cutting review, the extent of the evaluation purpose was multiple and the scope of the evaluation objects was broad.

Finally, in principle, it is clear that, in the five evaluation exercises, instrumental and conceptual utilisation of the evaluation findings was excellent. Therefore, the evaluation findings of the five R&D evaluation cases are expected to contribute to the future development of SET policies at the national level as well as to the immediate improvement of the related S&T or R&D programmes at the ministry level. However, there is a need for improvement in relation to the dissemination of the evaluation report of

some evaluation exercises. In particular, it is not reasonable that the evaluation reports of the BBSRC's two evaluation exercises were circulated only in the office, and was not made available to the public.

In conclusion, it is apparent that, among the five UK R&D evaluation systems, both the 2001 quinquennial review of six grant awarding research councils and the evaluation of funding for UK civil space activities are better examples of evaluation exercises in the UK. Meanwhile, the two evaluation exercises conducted by the BBSRC, namely the 2001 IAE and the evaluation of the GRF programme, could be considered as relatively good cases. On the other hand, it seems that there is more need for improvement in the 2000 cross-cutting review of science and research.

### **6.3. Meta-evaluation of the Korean National R&D Evaluation System**

Following the meta-evaluation of the UK R&D evaluation exercises, I shall now conduct the meta-evaluation of the five Korean R&D evaluation exercises, using the same model and a similar descriptive structure as for the UK case. In particular, I shall examine the extent to which the current state of the individual Korean R&D evaluation systems equates to the ideal state of an evaluation system from the perspective of policy evaluation theory. As ready noted, in order to conduct the meta-evaluation, use was made of the results of the five case studies of R&D evaluation systems described in Chapter 5, which were implemented at a different level.

#### ***6.3.1. Meta-evaluation of the Evaluation of Major Policy Issues of the MOST in 2001***

In early 2002, the OPM/the OGPC conducted an evaluation of the performance of central and local government in 2001. As a part of the evaluation of central administrative offices, the OGPC and the PEC evaluated 63 major policy issues in central government. In line with this work, the OGPC also evaluated some policy issues that are relevant to S&T and R&D affairs in the related ministries including the MOST, the MOCIE and the MIC. Of these, the evaluation of one major MOST policy issue, namely "the strategic development of leading edge technologies", is directly related to S&T and R&D policies at the national

level.

#### **6.3.1.1. Evaluation Paradigm**

(1) Evaluation Purpose : In general, the evaluation of government performance is conducted for the three following purposes : (1) to improve the effectiveness of policy formulation and its implementation in each individual ministry, (2) to ensure the accountability of the pertinent ministry for them and (3) to enhance the credibility of the government policies in the eyes of the citizens (OPM, 2001a). From these remits of the evaluation of government performance, it is evident that, in the evaluation of the major policy issue of the MOST in 2001, both “accountability” and “policy improvement” were considered as main evaluation purposes (OGPC and PEC, 2002b : 1, 4).

In evaluating the purposes of the evaluation of the MOST’s major policy issue, it is necessary to discuss both its current state and what would be a more desirable future system. Currently, it seems that the evaluation purposes proposed were relatively proper in appearance, in that the evaluation of the government performance is conducted on a six monthly basis. It is inevitable that the evaluation exercises conducted at short intervals will focus more on the direct and immediate evaluation purposes, including accountability and policy or programme improvement, than on more comprehensive evaluation purposes such as knowledge enhancement. However, given that the system and its actual exercise of the evaluation of the government performance need to be improved, it is clear that the evaluation purpose was not rational, because too much focus was placed on the short-term perspective. Therefore, as well as extending the intervals at which evaluations are held and adjusting the evaluation objects, there is a great need for the evaluation purposes to include knowledge enhancement, which is regarded as one of the long-term purposes of evaluation.

(2) Evaluation Object : In the evaluation of the MOST’s major policy issues, “the strategic development of leading edge technologies” was selected as one of the main evaluation objects. This main object comprised four sub programmes and 17 projects (OGPC and PEC, 2002b : 2-3). Therefore, it appears that the level of the evaluation object included both “sub programmes” and their “main projects”. On the other hand, the evaluation object was confined in scope only to the MOST’s activities. Thus, it is evident

that the evaluation object included specific sub programmes and more detailed projects that were related to the policy of the MOST.

From this, it is clear that the evaluation objects were inadequately set both in terms of level and scope. Therefore, the level of the evaluation object needs to be adjusted to cover the policy itself, in order to completely realise the initial intention of judging the performance of a certain policy, and not such that of sub programmes and detailed projects. Also, the scope of the evaluation objects has to be properly expanded to include related policies and programmes which have been carried out by many central ministries, including the MOE, the MOCIE, the MIC and the MOHW, in order to support and develop the leading edge technologies

(3) Evaluation Type : The evaluation of the government performance is basically conducted twice a year. While in the first half of the year, the rationality of policy formulation and the effectiveness of its implementation are examined, in the second half, the achievement of the policy objectives and their effect on the public are vigorously investigated (OGPC and PEC, 2001b : 6). Hence, this evaluation is explicitly intended to include an “interim” evaluation, an “output” evaluation and, where appropriate, an “impact” evaluation.

As a result, it seems currently that the types of the evaluation of government performance, including the evaluation of the major policy issues of the MOST, were suitable for the evaluation purposes. However, there is a doubt about how well the impact evaluation was conducted, bearing in mind that this evaluation exercise has been carried out twice a year. This is because the impact evaluation normally deals with long-term and indirect influences of the pertinent policy on the economy and society as a whole. Therefore, there is a great need to improve the focus of the evaluation type, changing it from a current interim evaluation to an output and impact evaluation, in accordance with the proposed improvement of both evaluation purposes and the evaluation objects.

Consequently, it is clear that, in this evaluation exercise, all three components of the evaluation paradigm were not rationally and suitably established, due mainly to the fact that the evaluation was conducted as a six monthly exercise. Therefore, to develop the evaluation of the government performance, firstly, there is a strong need to extend the

interval between evaluations, for example, so that they are conducted every two or three years rather than on a six monthly basis. Then, the evaluation paradigm, in particular the evaluation object and types, should be adjusted and amended to take advantage of the improvement of the evaluation timing.

### **6.3.1.2. Evaluation Resources**

(1) Evaluation Manpower : In general, three groups of manpower are involved in the evaluation of government performance, Firstly, the officials of the “Assistant Minister’s Office (AMO) for Policy Analysis and Evaluation” under the OGPC prepares the master plan, supports the PEC in conducting actual evaluation and coordinates evaluation results. As at April 2002, the total number of evaluation-related staff in this office was 35. Secondly, the members of the “PEC” conduct the actual evaluation in each area under their responsibility, with the assistance of experts from the support team. The PEC is normally composed of 30 members, including the Minister of the OGPC and 29 non-government experts, representing academia, industry, the media and other private organisations. Thirdly, the “support team”, whose mission is to assist the evaluation activities of the PEC with high expertise in each area, comprises 29 experts mainly from academia and related research institutes (PEC, 2002b).

Meanwhile, a total of four people were involved in the evaluation of the major policy issues of the MOST : one member of the PEC, one expert from the support team and two officials of the Assistant Minister’s Office (AMO) for Policy Analysis and Evaluation (OGPC and PEC, 2002b). Therefore, this evaluation was conducted through collaborative work between the staff of the AMO for Policy Analysis and Evaluation, and members of the PEC and its support team.

As a result, it is apparent that the external manpower, in particular, the members of the PEC, has actual responsibility for this evaluation, even though it receives administrative and, sometimes, professional assistance from the staff of the AMO for Policy Analysis and Evaluation as well as from the support team.

(2) Evaluation Unit : In general, the AMO for Policy Analysis and Evaluation under the



OGPC has the mission of an evaluation unit. Of the three director-generals in this office, the “Director-General I for Policy Analysis and Evaluation (PAE)” is responsible for the evaluation of non economic policies performed by 22 central administrative offices and consists of 21 members of staff (OPM, 2002a ; PEC, 2002a). Therefore, the Director-General I under the AMO for Policy Analysis and Evaluation acted as the evaluation unit in the evaluation of the strategic development of the leading edge technologies of the MOST. As a result, it appears that the assignment of the Director-General I to the evaluation unit was suitable, because it has independence and professionalism in managing the evaluation exercise.

(3) Evaluation Information : The evaluation information was drawn from three kinds of sources : performance reports, self-evaluation reports and extensive consultation. The performance report of individual ministries is expected to contain (1) an outline of policy, (2) the background to policy formulation and implementation (3) policy objectives and predictable effects and (4) the operating plan for sub programmes (OGPC and PEC, 2001b : 27-28). Meanwhile, the central administrative offices are required to must conduct a self-evaluation of major policy issues adopting the same evaluation methods and criteria as those used in the PEC’s exercise and to submit the self-evaluation report to the OGPC in advance of the evaluation work of the PEC (OGPC and PEC, 2001b : 33). Information was also obtained from extensive consultation with experts and the members of staff involved in the pertinent policies (OGPC and PEC, 2001b : 4-5 ; OGPC and PEC, 2002a : 4). Thus, we may see that much more use was made of qualitative information than of quantitative information in conducting the evaluation of the major policy issue of the MOST. In addition, it is clear that much more use was made of second hand information, for example, from performance reports and self-evaluation reports than of the first hand information such as that gained from extensive consultation, although both were utilised.

Consequently, it is believed that, bearing in mind the current state of the evaluation system of government performance, the overall evaluation resources were invested in a sufficient and suitable way. One distinction can be made in relation to the evaluation information. That is that the evaluator relied mainly on the self-evaluation reports of each ministry in order to obtain the information necessary for this evaluation. It seems that this had a close connection with the evaluation method, namely meta-evaluation approach,

discussed in the following sub-section.

### **6.3.1.3. *Evaluation Performance***

(1) Evaluator : In general, the evaluation of government performance is conducted by the PEC, with administrative assistance from the Assistant Minister's Office. Under the PEC, there are five sub committees that cover policy areas grouped into similar types. Each member of the sub committee has a responsibility for conducting the actual evaluation of major policy issues of one or two central ministries falling within his/her remit. He/she also receives professional assistance from the "support team" which is equivalent to a sub committee (PEC, 2002b).

One member, who belonged to the sub Committee for Economic Policies II, led the evaluation of the major policy issues of the MOST, with the assistance of one expert from the support team and two members of staff under the Director-General I for PAE (OGPC and PEC, 2002b).

From this, it is evident that the external evaluators led the exercise of the evaluation of the government performance, although internal evaluators provided them with essential assistance. Meanwhile, all external evaluators came from the private sector including academia, industry, the research institutes and the media. This is one way in which the Korean evaluation exercises were distinct from the UK evaluation cases.

(2) Evaluation Method : The members of the PEC firstly reviewed performance reports and self-evaluation reports. Next, in order to confirm the contents of these, they visited each ministry or site and also conducted an intensive consultation with both the experts concerned and related staff (OGPC and PEC 2001b : 5). Therefore, a typical qualitative method was adopted as the main evaluation method in the evaluation of major policy issues.

Meanwhile, individual sub programmes were given a certain grade on a four-tier grading system of very good, good, fair and unsatisfactory in a qualitative way, on the basis of the evaluation results of detailed projects. Then, these results were integrated at the policy level to decide how well the pertinent policy had been implemented and progressed by

comparison with the initial plan (OGPC and PEC, 2001a : 25-30).

From this, two distinct characteristics of the evaluation methods are evident : the adoption of the meta-evaluative approach and the application of the grading system. In particular, unlike in other Korean evaluation exercises, the meta-evaluative approach was adopted in order to judge and examine both performance reports and the self-evaluation reports utilised as the main sources of evaluation information. On the other hand, as in other evaluation exercises, the grading system was applied for the qualitative measurement of each sub programme.

(3) Evaluation Criteria : In order to examine the overall policy process, including policy formulation, policy implementation and policy achievement, the following six common evaluation criteria were adopted : (1) the validity of policy objectives, (2) the integrity of contents in policy objectives, (3) the efficiency of implementation, (4) the appropriateness of implementation, (5) the rate of achievement of objectives and (6) policy effectiveness. Meanwhile, in principle, individual ministries were asked to adopt these common evaluation criteria for their self-evaluation, even though they were allowed to choose some extra evaluation criteria (OGPC and PEC, 2001b : 6 ; PEC, 2002a). Thus, the evaluation criteria adopted were mainly qualitative in nature. However, there was no clear identification of the evaluation indicators in the related documents of the OGPC and the PEC.

In consequence, for similar reason as in the meta-evaluation of the evaluation resources, it is thought that the overall evaluation performance was relatively effectively implemented. However, there were some distinctive features of the evaluation performance. These include the professional background of the external evaluators as well as the introduction of the meta-evaluative approach and the adoption of the grading system in relation to the evaluation methods.

#### **6.3.1.4. Evaluation Utilisation**

(1) Evaluation Report : In general, the final report of the evaluation of government performance is open to the public by posting it on the website of both the OPM the PEC. It is also required to be submitted to the National Assembly. The final report comprises a

comprehensive report and many individual reports for each ministry (OGPC and PEC, 2001a : 23 ; OGPC and PEC, 2002a ; PEC, 2002a). The evaluation report on the major policy issues of the MOST included individual evaluation findings of the five sub programmes based on the six evaluation criteria (OGPC and PEC, 2002b : 9-24).

From this, it is evident that both the many individual evaluation reports and the comprehensive report focused more on the results of the evaluation exercise. For this reason, it seems that they lacked some important components of the evaluation, including an evaluation purpose, evaluation types, the evaluation information utilised, the evaluation methods adopted, an evaluation process and so on, which are normally expected to be included in an evaluation report. Therefore, it is clear that the detailed contents of the evaluation report need to be diversified to include the essential components of the evaluation exercises.

(2) Utilisation Type : In general, the evaluation findings of government performance are utilised for the following three purposes : (1) giving awards to excellent institutions and staff, (2) improving policy through the implementation of corrective measures and the conduct of an audit and (3) adjusting budget allocation for the following year (OPM, 2001a ; PEC, 2002a).

In line with this policy, in the 2001 exercise, four central administrative offices were given awards (OGPC and PEC, 2002a : 45). Also, a total of 323 projects in the area of major policy issues were identified as needing corrective measures. In the case of the MOST's major policy issue, a total of five out of 17 detailed policy projects were identified for corrective measures. These were required to be reflected in the operating plans for the following year (OGPC and PEC, 2002a : 44 ; OGPC and PEC, 2002c : 46-47). In addition, the OGPC notified the MPB and the ministries about the projects, which need to be subject to budgetary measures based on the evaluation findings (PEC, 2002a).

Therefore, it is apparent that the findings of the evaluation of government performance were largely utilised for short-term policy improvements, rather than long-term policy development. In particular, the findings of the evaluation of major policy issues of the MOST were considered as basic and important information which could be used to

identify problems at the stage of both policy formulation and implementation as well as to suggest corrective measures in the future to improve sub programmes and detailed projects of the pertinent policy. As a result, it is clear that the instrumental utilisation was emphasised rather than the conceptual utilisation, and this has a close connection with the evaluation paradigm, which is defined in terms of a somewhat narrow and short-term perspective.

In sum, it appears that in the evaluation of government performance, both the production of the evaluation report and its utilisation type need to be improved. In particular, the utilisation type of the evaluation findings should be extended to reflect the improved and amended evaluation paradigm, as already discussed in the meta-evaluation of the evaluation paradigm.

#### **6.3.1.5. *Synthesis***

In the evaluation of government performance and, in particular, the major policy issue of the MOST, there are essential requirements for the improvement of some components of the evaluation exercise. As already discussed, it is judged that these deficiencies resulted from the evaluation timing set up on a six monthly basis. As a result, the current evaluation system and its exercise have more features of annual monitoring rather than of *ex post* evaluation. Therefore, in order to achieve the initial objectives of the introduction of the evaluation of government performance, it is essential that the intervals between each evaluation exercise should be extended, for example, to two or three years.

It is clear that, when the above changes have been made, the overall evaluation paradigm also needs to be improved. One of its components, the evaluation objects, should be re-established at a more long-term strategic level rather than at short-term operating level. In particular, the scope of the evaluation objects in evaluations of major policy issues in the central ministries needs to be extended to cover related cross-departmental policies. For the same reason as indicated in the general discussion in relation to evaluation objects, the evaluation objects of the S&T and R&D policies should be adjusted in order to deal with as many related policies in the central administrative offices as possible, including the MOE, the MOCIE, the MIC and MOHW.

Then, both the evaluation purposes and the evaluation types also need to be adjusted and expanded in accordance with the rearrangement of the evaluation objects. Providing that the evaluation paradigm can be re-established at the level of policy evaluation itself rather than programme evaluation, it is expected that the improvement of the utilisation types of evaluation findings will follow.

### ***6.3.2. Meta-evaluation of the Evaluation of National R&D Programmes in 2001***

In early 2001, the NSTC conducted the evaluation of 161 national R&D programmes that had been implemented by 18 central ministries and administrative offices in the year 2000, as a part of that year's analysis, evaluation and pre-budget review of national R&D programmes. This evaluation comprised two types of work that were closely connected, although their aims were slightly different : the "programme evaluation" to judge the performance of R&D programmes and the "project review" to investigate the overlapping among R&D projects.

#### ***6.3.2.1. Evaluation Paradigm***

(1) Evaluation Purpose : In general, the analysis and evaluation of national R&D programmes mainly aims to improve the effectiveness of R&D investment funded by the government. This principal purpose includes the following sub purposes : (1) the "accountability" of R&D related organisations, (2) the "improvement" of national R&D programmes and (3) the "knowledge sharing" for the effective R&D management in the public sector (MOST, 2000b : 2 ; NSTC, 2001a : 2).

As in other Korean evaluation cases, this evaluation was conducted at relatively short intervals, namely every year, and more emphasis was placed on the effectiveness of the R&D investment than on other factors. As a result, currently it seems that the evaluation purposes of the national R&D programmes were relatively adequately proposed. However, in evaluating the evaluation purposes from a longer and broader perspective, it is apparent that the current evaluation purposes need to be adjusted to reflect the extension of the evaluation timing. This is because, in the large-scale R&D programmes,

there has been a tendency for their whole effect or influence to be produced in some period after their completion (Georghiou, 1998 : 40). Therefore, for the development of the evaluation system of the national R&D programmes, it is essential that the evaluation purposes will be strongly expected to include the knowledge enhancement regarded as one of the long-term evaluation purpose, alongside the extension of the evaluation timing and the adjustment of the evaluation objects.

(2) Evaluation Object : In principle, the evaluation object included all R&D programmes supported by the national budgets and public funds. However, some programmes that were not directly relevant to actual R&D activities were excluded from this exercise. As a result, 161 R&D programmes were selected as the detailed evaluation objects (MOST, 2000b : 6 ; KISTEP, 2002b : 94 ; Yang, 2002 : 35). Therefore, the scope of the evaluation objects contained most R&D programmes at the national level.

The actual level of the evaluation objects was differently defined depending on the programme evaluation or project review. In the programme evaluation, the detailed evaluation objects were 161 sub-programmes and these were categorised into four main groups and 12 sub groups (MOST, 2000b : 8 ; NSTC, 2001a : 12). On the other hand, in the project review, the evaluation objects were the 9,533 sub-projects that were the lowest units of the pertinent programmes (MOST, 2000b : 8 ; NSTC, 2001a : 13).

From this, it is evident that the 2001 NSTC evaluation exercise was actually conducted at the level of sub programmes or sub projects, even though its initial intention was to examine the overall performance of the programmes themselves. Therefore, there is a strong need for the level of the evaluation objects to be upgraded to that of the programme itself, given that the evaluation purposes would be re-established to cover a broader and longer perspective, for example, knowledge enhancement.

(3) Evaluation Type : In order to achieve the principal evaluation purpose, two different evaluations were conducted. One was to judge the performance of diverse R&D programmes. The other was to review the overlap among R&D projects and to identify linkages among them (NSTC, 2002a). Therefore, it is apparent that the NSTC's evaluation took a short term perspective and focused a great deal on both "interim" and "output" evaluation. Thus far, no impact evaluation has been conducted by the NSTC,

even though it is considered one of the main purposes of the evaluation of the national R&D programme. It is clear that this situation was closely connected with the evaluation purposes and the evaluation objects discussed above.

Consequently, it is clear that all three components of the evaluation paradigm were not proposed in rational and suitable terms. This is because this evaluation was initially introduced to enhance the effectiveness of the R&D investment, which is regarded as a short-term evaluation purpose. Therefore, for the future development of the evaluation system of the national R&D programme in the long-term, first of all, there is a strong need to upgrade the evaluation objects, from sub programmes and sub projects to programmes themselves. Then, the evaluation purposes need to be adjusted and the evaluation types diversified.

#### **6.3.2.2. *Evaluation Resources***

(1) Evaluation Manpower : Three groups of manpower were involved in the evaluation of the NSTC : (1) members of the evaluation committees to carry out the actual evaluation, (2) internal members of staff in the MOST and the KISTEP to act as the secretariat of the NSTC and the evaluation committees, and (3) the supporting team to provide administrative assistance.

The evaluation committee was entirely made up of external experts and the following three kinds of sub committee were organised : (1) the steering committee, (2) the various evaluation committees for sub R&D programmes and (3) the various review committee for R&D projects. "The Steering Committee" comprised 19 external members and coordinated the evaluation results from 14 project review committees and eight programme evaluation committees. "The Evaluation Committee for Sub Programmes" comprised eight sub committees and a total of 83 experts were involved in these committees. "The Review Committee for R&D Projects" was also divided into 14 sub committees and a total of 140 experts were involved in these committees (NSTC, 2002a : 325-350). As a result, a total of 242 members sitting on the three types of committee were engaged in the evaluation of 161 sub R&D programmes and the review of 9,533 R&D projects.



In the MOST, “the Science and Technology Assessment Division” was responsible for the master plan of this evaluation and, as of April 2002, comprised eight members of staff (MOST, 2002b). On the other hand, in the KISTEP, “the Division of National Research and Development Evaluation” is in charge of the overall evaluation work at the operating level and, as of April 2002, consisted of 21 members of staff (KISTEP, 2002b). Thus, in the 2001 exercise, approximately 30 internal members of staff from the MOST and the KISTEP were involved. Meanwhile, the supporting team comprised 16 members, who were R&D management experts from central administrative offices and their affiliated institutions (NSTC, 2002a : 351).

As a result, around 270 members, including 240 external experts and 30 internal staff, were fully or partially engaged in the evaluation of national R&D programmes in 2001. One particular characteristic of the evaluation manpower is that the three evaluation committees were entirely made up of external experts from academia, industry, the research institutes and the media. Hence, unlike in some UK evaluation cases, the related members of staff in the government ministries did not participate in the actual evaluation exercise.

(2) Evaluation Unit : Both “the Science and Technology Assessment Division” of the MOST and “the Division of National Research and Development Evaluation” of the KISTEP acted as the evaluation units. While the role of the former was to set up the master plan for the analysis, evaluation and pre-budget review of national R&D programmes, the latter’s role was to implement this plan and to assist the evaluation activities of the committees (MOST, 2002b ; KISTEP, 2002b). Among 21 members in the Division of National Research and Development Evaluation of the KISTEP, four members of staff devoted themselves fully to this evaluation. As a result, two divisions acted together as the evaluation unit, even though they were also responsible for other matters in each organisation. Therefore, it was appropriate that the independent and permanent organisations of the both the MOST and the KISTEP were designated as evaluation units.

(3) Evaluation Information : Most of the important information came from two kinds of source : (1) various formal documents submitted by individual ministries and (2) presentations made by the programme managers. Each central administrative office had

to submit the required materials to the MOST, which had become the NSTC secretariat by the end of February. In addition, each programme manager was required to give a presentation of the research outputs in order to supplement the related materials (NSTC, 2001a : 2-3 ; Yang, 2002 : 18). As a result, qualitative information was utilised as the main source for this evaluation. Also, it is evident that this evaluation work relied mainly on the second hand information produced by the individual ministries to be evaluated, rather than on the first hand information collected by the evaluators themselves.

In sum, it seems that, at the current stage, the overall evaluation resources were invested in sufficient and proper manner. In particular, some distinctions can be made regarding both the evaluation manpower and the evaluation information. Firstly, the evaluation committees were entirely comprised of external experts from the private sector, including academia, industry, the research institutes and the media, who had no interest in the programmes to be evaluated, and did not include any staff from the government ministries who were regarded as having some interest in them. Secondly, in order to obtain the evaluation information, the evaluation committees utilised mainly the performance report submitted by individual ministries, rather than making efforts to get additional information actively.

#### ***6.3.2.3. Evaluation Performance***

(1) Evaluator : Of the three groups of evaluation manpower, it was the members of two types of evaluation committee, namely the evaluation committee for sub R&D programmes and the review committee for R&D projects who conducted the actual evaluation exercise. All the members of the two committees came from academia, the research institutes and industry. In particular, there were more members from academia than from industry and the research institutes, although their proportion was not much higher. So, it is evident that this evaluation was entirely conducted by external evaluators, even though there was much assistance from members of staff in the KISTEP and the supporting team.

(2) Evaluation Method : In order to judge the performance of R&D programmes, the evaluation committees reviewed official performance reports submitted in advance by individual central ministries and agencies (NSTC, 2001a : 14-15). Thus, a qualitative

analytical approach, in the form of a literature review, was the principal evaluation method adopted. Meanwhile, a five-tier grading system, consisting of excellent (A grade), good (B grade), fair (C grade), unsatisfactory (D grade) and need for improvement (E grade), was applied to the preparation of the final evaluation results in each programme.

(3) Evaluation Criteria and Indicator : In the programme evaluation, three main criteria were adopted : “the appropriateness of the programme objectives”, “the efficiency of implementation” and “the level of performance and outputs”. The following additional criteria were also used : “the need for adjusting the programme’s scope”, “integration and/or linkage with other programmes” and “re-planning of the programme”. The weights for the evaluation criteria differed according to the attributes of programme groups (NSTC, 2001a : 14-15 ; Yang, 2002 : 35). In the project review, “the redundancy of R&D projects” and “the need for integration and/or linkage with other projects” were adopted as the major evaluation criteria (NSTC, 2001a : 16). Thus, it appears that qualitative evaluation criteria were adopted to a large extent in the 2001 exercise. Meanwhile, there was no clear identification of the evaluation indicators in the related official documents.

Consequently, it seems, at the current stage, that the overall evaluation performance was implemented in a suitable manner for the realisation of the evaluation purposes. Also, as in other Korean cases, a grading system was applied to the integrated overall final evaluation results of each individual R&D programme. Therefore, it can be said that there were no differences in evaluation performance, compared to other general programme evaluation cases.

#### **6.3.2.4. Evaluation Utilisation**

(1) Evaluation Report : The evaluation report was published and has also been available on the website of both the NSTC and the MOST. Many parts of the report were devoted to the discussion of evaluation findings and major recommendations in relevant to individual sub R&D programmes and R&D projects (NSTC, 2001a). Thus, it is clear that the production and the dissemination of the evaluation report were good and its detailed contents were also appropriate.

(2) Utilisation Type : Two types of evaluation utilisation were made. Firstly, the evaluation findings were directly reflected in the allocation of research funds to national R&D programmes for 2002 through the process of the pre-budget review (MOST, 2001a ; NSTC, 2001 : 2). Next, seven types of recommendations for the improvement of national R&D programmes at national level were suggested. As a result, a total of 70 programmes were identified as needing improvement in 2002 (NSTC, 2001 : 99-100, 110 ; Yang, 2002 : 42, 45). On the other hand, two types of recommendation were made, concerning the improvement of R&D projects. Among the R&D projects, 43 were seen as redundant investment and 121 needed to be linked with other projects (NSTC, 2001a : 110 ; Yang, 2002 : 47).

Therefore, it can be said that the findings of the evaluation of national R&D programmes were utilised mainly for the short-term improvement of R&D programmes and projects, rather than the long-term development of R&D programmes themselves and the related S&T or R&D policies. In particular, among the diverse utilisations of the evaluation, more emphasis was placed on the allocation of research funds among national R&D programmes for the following year, which is recognised as one of the most powerful and effective policy means. It seems that these short term and instrumental types of utilisation of evaluation findings were closely connected with the overall evaluation paradigm, in particular the fact that the evaluation objects were fixed at a low level.

In consequence, it appears that the findings of the evaluation of the national R&D programmes were utilised in too narrow and restricted a way, admitting that the three components of the evaluation paradigm were established in the perspectives of the operational management for sub R&D programmes and projects. Hence, the focus of the evaluation utilisation needs to gradually transfer from the current instrumental utilisation to the conceptual utilisation.

#### **6.3.2.5. *Synthesis***

From the meta-evaluation of the 2001 national R&D programmes, it is evident that there are improvements required in some components of the evaluation exercise. As already discussed, these shortcomings resulted from the original intention to enhance the effectiveness of the R&D investment as well as to avoid redundancy among the R&D

projects. As a result, it is believed that the current system of the evaluation of national R&D programmes and its implementation have more of the characteristics of annual monitoring or interim evaluation than of *ex post* evaluation.

Therefore, in order to develop this evaluation exercise into a more meaningful policy means, it is essential that the evaluation objects should be upgraded to programme level from the current level of sub programmes and sub projects, alongside the extension of the interval between evaluations, for example, from one to two or three years.

It is evident that, once the above changes have been made, improvements in other components of the evaluation paradigm should follow. That is, both the evaluation purposes and the evaluation types also need to be adjusted and expanded in accordance with the rearrangement of the evaluation objects. Providing that the overall evaluation paradigm can be re-established at the level of R&D programme evaluation rather than R&D project evaluation, it is expected that the types of utilisation of the evaluation findings will include both instrumental and conceptual utilisation.

### ***6.3.3. Meta-evaluation of the Evaluation of Five Research Councils in 2001***

The OGPC evaluated the performances and plans of five research councils implemented in 2000, following by the first exercise in 1999. The OGPC is responsible for the supervision of research councils and their member research institutes, which are the principal actors in public R&D activities including those in the social sciences.

#### ***6.3.3.1. Evaluation Paradigm***

(1) Evaluation Purpose : This evaluation was conducted with the intention of both maximising the effects of the research councils' activities and ensuring the legitimacy and the effectiveness of their management (OGPC, 2001a : 2 ; ECRC and KISTEP, 2001 : 4). From this, it is apparent that the OGPC conducted this evaluation in order to secure both the "accountability" and the "improvement" of the management of research councils. Therefore, it is evident that the evaluation purposes of the five research councils placed

much more emphasis on the current managerial matters of the research council system than on the strategic affairs of R&D development in the public sector.

(2) Evaluation Object : In the evaluation of the five research councils, two main evaluation objects were adopted : (1) the research council's own activities and the mission of supporting their member GRIs, (2) their internal management systems. Each evaluation object was given a different weight according to its attributes : 70% of the weight fell on the first evaluation object and 30% on the second one. The two evaluation objects were again divided into more specific sub objects, reflecting the evaluation purposes (OPM, 2001c ; ECRC and KISTEP, 2001 : 5-7).

Thus, it appears that this evaluation focused mainly on managerial matters at the operating level. In particular, more weight was placed on the effectiveness of the mission in supporting member research institutes than on the appropriateness of their internal management system.

(3) Evaluation Type : In this evaluation, firstly, an "output" evaluation was conducted in order to judge the past performance of the internal and external management systems. In addition, the "assessment" was carried out to investigate prospective the R&D planning and the development plan for member institutes (ECRC and KISTEP, 2001 : 5-7). Therefore, it seems currently that appropriate evaluation types were adopted to correspond to the initial evaluation purposes discussed above.

Thus, it is apparent that the overall evaluation paradigm focused on the managerial matters of the research council system at the operating level. This is because the newly introduced research council system has been considered slightly unstable and, sometimes, unfamiliar, and subsequently the Korean government has made many efforts to settle it down as soon as possible. Accordingly, it is judged that the current evaluation paradigm was established in a rational and appropriate way. However, if the government policies in relation to the research councils are transferred from the managerial matters to the long-term and strategic issues, the evaluation paradigm should also be rearranged to reflect the change of government policies.

### **6.3.3.2. Evaluation Resources**

(1) Evaluation Manpower : Two groups of manpower were involved in this evaluation : one was the external manpower from the evaluation committee and the KISTEP, whereas other was the internal manpower from the “Director-General for R&D Support” under the “Assistant Minister’s Office for Regulatory Reform” in the OGPC (OGPC, 2001a ; ECRC and KISTEP, 2001).

There were two types of external manpower engaged in this evaluation. Firstly, the evaluation committee comprised 10 members, all of whom came from the private sector, including academia, industry, the research institutes and the media (ECRC and KISTEP, 2001 : 144). Next, two members of the “the Division of National R&D Evaluation” of the KISTEP acted as the secretariat to the evaluation committee, on the basis of the contract commissioned by the OGPC (ECRC and KISTEP, 2001 : 10 ; KISTEP, 2002b). Meanwhile, the six internal members of staff of the Director-General for R&D Support prepared the master plan and circulated the evaluation manual to the research councils (OPM, 2002a). As a result, around 15 people were involved in the evaluation of the five research councils in 2001.

In consequence, as in other Korean cases, there were strict differences in the roles of the external and the internal evaluation manpower, in that the former conducted the actual evaluation and acted as the secretariat to the committee, whereas the latter prepared the master plan and produced the evaluation manual.

(2) Evaluation Unit : The Director-General for R&D Support under the OGPC was the evaluation unit that was responsible for managing the overall evaluation exercise. In particular, both “the Support Team for Humanities, Economic and Social Sciences Research” and “the Support Team for Science and Technology Research”, had a principal duty of planning and supervising this evaluation exercise. There were three members of staff in each team (ECRC and KISTEP, 2001a : 10 ; OPM, 2002a). Therefore, it seems that the designation of two support teams to the evaluation units was suitable, because they have a responsibility for the overall management of the research councils and their member research institutes.

(3) Evaluation Information : The main evaluation information was extracted from both the materials submitted by each research council and the responses to the questionnaire survey conducted by the committee. Firstly, the individual research councils were required to submit to the evaluation committee a total of 33 types of materials, which were slightly different according to each sub evaluation object (OGPC, 2001a : 8-33). Secondly, the questionnaire survey also provided the committee with useful information. It covered five main categories and contained 10 questions that corresponded to some important evaluation objects (ECRC and KISTEP, 2001a : 125-136). From this, in general, more emphasis was given to qualitative than to quantitative information, even though both might be required, depending on the attributes of the evaluation object. Meanwhile, unlike in some Korean evaluation exercises which relied mainly on second hand information submitted by the organisations to be evaluated, first hand information collected by the evaluators themselves was also utilised.

In conclusion, it can be said currently that the overall inputs of evaluation resources was sufficient and proper. In particular, some distinctions can be drawn in relation to both the evaluation manpower and the evaluation information. Firstly, as in other evaluation cases, the evaluation committees were entirely made up of external experts from the private sector, namely academia, industry, the research institutes and the media, who have high expertise in relation to both the management and research matters of the research councils. Secondly, the first hand information collected by the evaluation committee through conducting the questionnaire survey was adopted as another important source for the evaluation, even though it did not have the same weight.

#### **6.3.3.3. *Evaluation Performance***

(1) Evaluator : The actual responsibility for this evaluation was held by the “Evaluation Committee for Research Councils”. It was chaired by a private sector professional and included 10 members from academia, industry, the research institutes and the media (ECRC and KISTEP, 2001 : 144). The committee was divided into four sub committees that were in charge of detailed sub evaluation objects (ECRC and KISTEP, 2001 : 10-11). Therefore, it is apparent that only external evaluators participated in the actual evaluation exercise, even though they obtained a great deal of administrative assistance from the internal staff of the KISTEP.



(2) Evaluation Method : The main evaluation methods adopted were a literature review, a site visit and a questionnaire survey. Each sub committee reviewed diverse materials submitted by individual research councils as well as the results of the questionnaire survey. They also visited all research councils to get supplementary information (ECRC and KISTEP, 2002 : 9-11). In addition, the committee conducted a questionnaire survey. A total of 200 questionnaires were distributed to 30 to 60 people related to each research council. The results of the questionnaire were analysed in both a quantitative and qualitative way and were also utilised as reference information for the evaluation exercise (ECRC and KISTEP, 2001 : 125-135). As a result, it seems that the adoption of a mainly qualitative evaluation method was directly related to both the evaluation purpose and the evaluation object, which were defined and set up in a way that was more qualitative than quantitative.

(3) Evaluation Criteria and Indicator : Two or three evaluation criteria were firstly adopted to judge each sub evaluation object. Then, each evaluation criterion was again divided into one to three evaluation indicators. As a result, a total of 21 evaluation criteria and 36 evaluation indicators were adopted (ECRC and KISTEP, 2001 : 6-7). For each indicator, the evaluation was conducted in a qualitative way and then a certain grade was given on a five-tier grading system (ECRC and KISTEP, 2001 : 8). Therefore, it appears that qualitative evaluation criteria and indicators were mainly adopted, although a few quantitative ones were utilised depending on the evaluation object.

Consequently, it is thought that the overall evaluation performance was properly implemented in accordance with the evaluation purposes. In fact, only one distinction can be drawn regarding the evaluation method. Unlike in other evaluation cases, in this evaluation, the evaluation committee conducted the questionnaire survey with the intention of investigating the related people's opinions about the performance of R&D management by the research councils. Therefore, the adoption of the questionnaire survey was considered as one of the efforts of the OGPC to diversify the evaluation methods.

#### **6.3.3.4. Evaluation Utilisation**

(1) Evaluation Report : Although the evaluation report was published, it has not been

available on the website of related institutions. In particular, in the part relating to the evaluation results for each research council, it provided an overview of each council, a synthesis of the evaluation findings, the results of the evaluation of individual indicators and an account of the correction of unsatisfactory matters identified in the previous year's evaluation (ECRC and KISTEP, 2001). As a result, there is a need to improve the dissemination of the evaluation report, even though its detailed contents are relatively good.

(2) Utilisation Type : There was no explicit announcement about the detailed utilisation of evaluation findings in either an evaluation manual or an evaluation report. However, according to the statutes concerned, the evaluation findings are to be notified to the minister of the MPB. In addition, those in relation to the three research councils in the S&T sector, namely the KRCF, the KORP and the KOCI, are to be also notified to the NSTC (OPM, 2001c ; OGPC, 2001a : 1). It is expected that the evaluation findings were mainly utilised as policy tools to adjust the budget for R&D activities in the public sector. Therefore, it appears that the utilisation of the evaluation findings focused on the short term and instrumental perspective.

In consequence, there is a need for improvement in relation to the utilisation types of evaluation findings. That is, the utilisation type of the evaluation findings should be expanded to cover the long term and conceptual perspective, provided that the evaluation paradigm is re-established to reflect the change of government policy in relation to the management of the research councils from managerial matters to long-term and strategic issues.

#### **6.3.3.5. *Synthesis***

Having evaluated the evaluation exercise of the five research councils, it seems that, in general, the overall status of the evaluation system and its exercise is that they are at an early stage of development. As a result, there is more need to improve the evaluation system itself and the actual implementation of most components of the related evaluation practice than in other evaluation cases. It is believed that these deficiencies came not only from a relatively shorter history of involvement with and experience of the research council system but also from government policies that place more emphasis on the

managerial matters of the research councils themselves than on systematic and strategic R&D governance structure in the public sector.

Therefore, in order for this evaluation exercise to be developed into a useful policy means in the public R&D management system at national level, there is a strong need for the government policies concerning the research council system to be changed so that they have a coordinative and supporting perspective rather than the current managerial one.

It is clear that, when the above changes have been made, both the re-design of the overall evaluation system and improvements of the major components of the actual exercises should also be pursued. In particular, it is very important that the intervals between evaluation exercises be extended and that the evaluation purposes be set out in broader terms. Providing that both the evaluation timing and the evaluation purposes can be improved in accordance with the change of the government policies concerned, other evaluation components, including the evaluation resources, the evaluation performance and the evaluation utilisation, will become more practical and instrumental matters at the operating level.

#### ***6.3.4. Meta-evaluation of the Evaluation of Member Research Institutes by the KOCI in 2002***

The five Korean research councils have carried out the evaluation of their member research institutes, namely the GRIs, every year since 2000. In early 2002, as a part of this collective work, the Korea Research Council for Industrial Science and Technology (KOCI) conducted the evaluation of the seven member research institutes' performance in 2001 and their research plans for 2002.

##### ***6.3.4.1. Evaluation Paradigm***

(1) Evaluation Purpose : The evaluation purposes were clearly announced in the official documents including the evaluation manual and the evaluation report (KOCI, 2002b ; (KOCI, 2002c : 1). From these statements, it is clear that the "improvement" of the management system and research programmes of member institutes was the most

important evaluation purpose. In addition, both “accountability” for the overall management in each research institute and the “dissemination of knowledge” to related research fields were considered as further evaluation purposes. Therefore, it is appropriate that more emphasis was placed on the improvement of member research institute rather than on other evaluation purposes, bearing in mind that the new S&T governance structure, namely the research council system, was recently introduced in the public sector R&D management system.

(2) Evaluation Object : The evaluation object comprised two main areas : R&D matters and internal managerial matters. Each main evaluation object was given the same weight and was again divided into several sub evaluation objects, which were given a different weight, depending on the characteristics of individual research institutes. As a result, a total of 10 specific sub evaluation objects were selected (OPM, 2001c ; KOCl, 2002c : 9-11).

Therefore, it seems that the level of evaluation object dealt with is operational matters observed from a short-term perspective, whereas the scope of the evaluation object covered the past performance of R&D activities and internal management, although some sub evaluation objects included past performance as well as future plans where appropriate. It is evident that the setting of the evaluation object at managerial level was closely connected with the evaluation purposes discussed above.

(3) Evaluation Type : In this evaluation, much emphasis was placed on the investigation of the overall past performance of each institute. However, a part of the judgement was made in relation to the future plans and strategies of the institute management where appropriate (KOCl, 2002c : 9-11). From this, it is evident that the KOCl work was carried out both as a retrospective “output” evaluation and as a prospective “assessment” of research plans. It is inevitable that the main type of this evaluation was an output evaluation, in that the evaluation purposes focused on the improvement and the accountability of the member research institutes.

In consequence, it seems that the overall evaluation paradigm was relatively rationally and adequately proposed, because the Korean government had a principal policy to improve the efficiency and the effectiveness of the management system for the GRIs by

conducting this evaluation exercise. However, if the government missions in relation to the GRIs are changed from acting as the administrative manager in the public sector R&D activities to playing the role of supporter and coordinator, the evaluation paradigm should also be revised to reflect the change in government policies. In particular, the evaluation purposes should be extended to include knowledge enhancement. This is because, in Korea, the GRIs have played and will be expected to play a principal role in public sector R&D activities. Meanwhile, regardless of the change in government policies, it seems that the numbers of sub evaluation objects currently need to be reduced, because too many evaluation objects will place on the administrative burden on both the evaluator and each research institute to be evaluated (Lee, 2001b ; Lee, 2002b).

#### **6.3.4.2. Evaluation Resources**

(1) Evaluation Manpower : Two groups of manpower were involved in the KOCI's evaluation : the external experts making up the evaluation committee and the support team, and the internal members of the KOCI. "The Committee for the Evaluation of Research Institutes" comprised 17 external experts : 12 members of the Planning and Evaluation Committee of the KOCI<sup>4</sup> and five specialists from outside. "The support team", chosen on the basis of open competition, comprised five specialists from the Science and Technology Policy Institute (STePI) (KOCI, 2002c : i, 12-13). Meanwhile, three members of staff in "the Evaluation and Management Team" made an evaluation plan and produced an evaluation manual after consultation from the OGPC (KOCI, 2002a).

From this, we may see that this evaluation was conducted through the collaborative work of the external experts and the internal members of staff. One particular characteristic of the external experts is that most of them came from the Planning and Evaluation Committee of the KOCI, which was entirely made up of private sector professionals in related S&T areas of the research council's responsibility. Therefore, it is believed that the external manpower, namely the members of the evaluation committee, has a relatively

---

<sup>4</sup> Initially, the Planning and Evaluation Committee, one of the two standing committees of the KOCI, comprised experts from academia, the research institutes and industry, whose major interest related to research areas of individual research institutes, such as oriental medicine, industrial technology, IT technology, food science, machinery and materials, chemistry and electrotechnology.

high expertise in the S&T areas of the member research institutes as well as in their internal managerial matters.

(2) Evaluation Unit : The evaluation unit was “the Evaluation and Management Team” who were responsible to the general secretary of the KOCI. Three members of this team were directly engaged in this evaluation (KOCI, 2002a). However, in order to ensure the expertise of the evaluation exercise, the KOCI organised “a support team” on the basis of a contract, although this team was not formally responsible for the evaluation exercise (KOCI, 2002b : 4 ; KOCI, 2002c : i ; STEPI, 2002a). Therefore, it is clear that certain limitations of the evaluation unit, for example the shortage of manpower and its relatively low expertise, were complemented by the outside support team.

(3) Evaluation Information : The main source of information was “the performance report” submitted by each member research institute. The evaluation manual gave detailed guideline for the preparation of the performance report (KOCI, 2002b ; 15-67). From this, it appears that both qualitative and quantitative information was adopted as the main source for evaluation, although the types and their weight varied slightly depending on the main and sub evaluation objects. Examples of the quantitative information included articles and proceedings published, various intellectual properties (IPs) applied for and registered, international collaborative research and so on.

Consequently, it seems that the overall evaluation resources were properly and suitably invested for the realisation of the current evaluation paradigm. As already noted, the evaluation manpower and the evaluation unit had some distinct characteristics. One of these was that most of the external evaluation manpower already had not a little knowledge about the two evaluation objects, namely R&D matters and internal management matters, through participating in the planning and evaluation committee. Another was that the support team was organised on a temporary basis in order to supplement the internal evaluation unit.

#### **6.3.4.3. *Evaluation Performance***

(1) Evaluator : To secure the expertise and the objectivity of the evaluation, the OGPC and the KOCI have a policy to conduct the evaluation of the GRIs through the

commission of evaluation exercises to outside experts or institutions (KOCI, 2002b : 4). In line with this policy, “the Committee for the Evaluation of Member Research Institutes”, consisting of 17 external specialists from the planning and evaluation committee as well as from outside, conducted the KOCI’s evaluation. Their major interests included related S&T areas, including oriental medicine, industrial technology, IT technology, food science, machinery and materials, chemistry and electrotechnology, as well as social sciences such as knowledge management, accounting, human resources and organisation theory (KOCI, 2002a ; KOCI, 2002c : 12-13). Therefore it is clear that the external evaluators had a relatively high expertise in both R&D matters and the internal management matters to be evaluated.

(2) Evaluation Method : The evaluation methods used were a literature review and a site visit. The committee firstly reviewed the performance report submitted by the individual research institutes. Then, it visited each institute and conducted interviews with some related members of staff to obtain supplementary information (KOCI, 2002b : 14 ; KOCI, 2002c : 7). Therefore, it appears that the key evaluation methods adopted were a literature review and a site visit including interviews, which are regarded as typical qualitative methods.

(3) Evaluation Criteria and Indicator : To examine 10 sub evaluation objects, one to four evaluation criteria were adopted to reflect the attributes of each evaluation object. Then, each evaluation criterion was again divided into two to nine detailed evaluation indicators. As a result, in the 2002 KOCI’s evaluation work, a total of 22 evaluation criteria and 45 evaluation indicators were adopted (KOCI, 2002c : 9-11). Meanwhile, a five-tier grading system was applied to this evaluation exercise (KOCI, 2002c : 8). As a result, it is said that evaluation criteria and indicators adopted were generally qualitative, although a few quantitative ones were utilised. In addition, it is clear that too many evaluation criteria and indicators were adopted, although this had something to do with the number of evaluation objects.

In sum, it is thought that the overall evaluation performance was relatively efficiently implemented, even though there is a need for improvement in the evaluation criteria and indicators. That is, the numbers of the evaluation criteria and indicators need to be reduced to a degree that will not cause administrative burdens on related people and

institutes, alongside the adjustment of the evaluation objects.

#### **6.3.4.4. Evaluation Utilisation**

(1) Evaluation Report : The evaluation report was published and has been available on the KOCI website. In particular, the section concerning the results of the evaluation of individual research institutes included a synthesis of the evaluation findings for each main evaluation object and recommendations for future development (KOCI, 2002c). Therefore, it is clear that the evaluation report was well prepared and broadly disseminated to related people and organisations.

(2) Utilisation Type : The final evaluation findings were submitted to the Prime Minister and the Minister of the MPB. Each member institute was required to reflect these findings in their 2002 strategies and plans for internal management and R&D programmes. The correction work suggested by the evaluation findings will also be regularly monitored by the KOCI (KOCI, 2002c : 3, 80). Hence, we may see that more emphasis was placed on the short-term improvement of the internal management system and R&D management at the level of each research institute, rather than on the mid and long-term development of R&D activities in the public sector at national level. It seems that these results had a relatively close connection with the evaluation purpose being defined at the operating level.

In consequence, it is clear that the utilisation of the evaluation should be adjusted to cover the conceptual perspective as well as the instrumental one, once the related government policies have been changed to deal with more strategic affairs in public R&D activities and the evaluation paradigm has also been revised in accordance with this change.

#### **6.3.4.5. Synthesis**

From the results of evaluating the evaluation exercise of the seven GRIs under the KOCI, it is evident that the current state of the evaluation system is at an early stage of development. Despite the early stage of the exercise, it seems that, in general, the 2002 evaluation was relatively properly and effectively conducted, in that it made a contribution to the government policy to induce member research institutes to reform



their internal management system actively in order to improve research productivity.

However, it is strongly recommended that the evaluation system of the GRIs and its exercise should be re-designed so that it can function as a meaningful policy means in the area of public sector R&D management. For this, first of all, the government missions in relation to the GRIs should be changed from acting as the administrative manager in public sector R&D activities to playing the role of supporter and coordinator. Then, it is essential that the interval between evaluations should be extended from a year to a longer period of, for example, three or four years. Since the Korean GRIs have played and will be expected to play a principal role in public sector R&D activities, there is a strong requirement for their R&D activities allocated to them on a more stable and predictable base.

It is necessary that, once the above changes have been made, the re-settlement of the overall evaluation system and improvements of the major components in the actual exercises should be pursued. In particular, it is very important that the evaluation purposes are set up in broader terms. Providing that both the evaluation timing and the evaluation purposes can be improved in accordance with the change of the government policies concerned, other evaluation components such as the evaluation resources, the evaluation performance and the evaluation utilisation, will become more practical and instrumental matters at the operating level.

#### ***6.3.5. Meta-evaluation of the Evaluation of the Creative Research Initiatives (CRI) by the MOST in 2000***

The MOST has a policy to evaluate all of its research projects and to use the findings to inform future funding decisions. In accordance with this policy, in 2000, for the first time, it evaluated the interim achievements of 27 R&D projects under the Creative Research Initiatives (CRI) that had been started in 1997. The aim of this programme is to strengthen the national potential for technology competitiveness through creative basic research. The project leader, who is selected on the basis of strict criteria based on creativity, leadership, research experience and so on, has exclusive authority and responsibility in managing the project (MOST, 2002c).

#### **6.3.5.1. Evaluation Paradigm**

(1) Evaluation Purpose : This evaluation was conducted in order mainly to decide whether 27 R&D projects started in 1997 would be given further research funding for the next stage. In addition, the MOST conducted this evaluation to maintain the original concepts and framework of the CRI scheme by making an appropriate response to the trends of technological change (MOST and KISTEP, 2000b : 1). Therefore, it is clear that this evaluation exercise concentrated on the “accountability” of the project leader and the “improvement” of the programme itself. So, it is clear that the evaluation purpose of the CRI programme was proposed very rationally in that the evaluation was conducted at the level of specific R&D projects.

(2) Evaluation Object : As at June 2000, this exercise included 27 R&D projects started in the year 1997 (MOST and KISTEP, 2000b : 2, 5). Therefore, it appears that the level of the evaluation object was the individual projects that made up the CRI programme. Meanwhile, the scope of the evaluation object was confined to existing projects, which had already produced some research outputs in the previous three years. As a result, it is believed that the level and the scope of the evaluation object were properly determined, bearing in mind the initial intention of this evaluation work that has been discussed.

(3) Evaluation Type : This evaluation mainly aimed to examine and judge what individual R&D projects had achieved in terms of research outputs and their quality, the capability of the project leader, future research plans and so on (MOST and KISTEP, 2000a : 12). Thus, it is apparent that this evaluation focused on an “output” evaluation of past performance and an “assessment” of future research strategy.

Consequently, in general, it appears that the three components of the evaluation paradigm were rationally and suitably established in order to achieve the initial objectives of the introduction of this evaluation.

#### **6.3.5.2. Evaluation Resources**

(1) Evaluation Manpower : Two groups of evaluation manpower were engaged in this evaluation : firstly many external experts such as (1) evaluation panels, (2) international

evaluators, (3) volunteer reviewers through the internet and (4) the Planning and Management Committee (PMC) for the CRI, and secondly the internal members of staff of “the Division of National R&D Planning and Management” in the KISTEP.

Firstly, “four sub evaluation panels” had a principal mission to conduct the evaluation through the integration of results from the international evaluators and volunteer reviewers. Each sub evaluation panel comprised around 12 members, and a total of 50 external experts were involved in the four sub evaluation panels (MOST and KISTEP, 2000b : 3). Secondly, three prominent international experts for each project also played a major role in the evaluation and a total of 77 international experts from 12 countries, including the USA, Japan and the UK, were involved in the evaluation of 26 R&D projects (MOST and KISTEP, 2000b : 3). Thirdly, for each project, around 30 “volunteer reviewers” from both home and overseas were invited to the evaluation through the internet. Their evaluation results were utilised as important reference information for the evaluation panels (MOST and KISTEP, 2000a : 9 ; MOST and KISTEP, 2002b : 6). Finally, “the Planning and Management Committee for the CRI”<sup>5</sup> was involved in the evaluation through the review and synthesis of results from other evaluators. (MOST and KISTEP, 2000b : 7, 12-13).

In addition, six staff from “the Programme Coordinator (PC) for Space Science and Fundamental Research” were partially involved in the evaluation through preparation of the evaluation plan and management of the detailed evaluation process (KISTEP, 2002b).

As a result, around 140 external experts from both Korea and overseas were directly engaged in the evaluation of 27 R&D projects under the CRI, and many scientists were also indirectly involved in this exercise through conducting the evaluation on the internet. From this, we may draw one feature of the evaluation manpower involved that is very different from that in other Korean evaluation exercises. That is, many prominent international scientists directly participated in the evaluation process. It is clear that this represented the MOST’s policy, which was intended to secure as much fairness and

---

<sup>5</sup> This committee was initially organised to establish the basic strategy for the CRI programme and then fully engaged in the overall management of that programme. The total number of members of the committee was 18. Of these, 16 were prominent scientists from academia, industry and the research institutes and two were related senior members of staff in the MOST and the KISTEP (MOST and KISTEP, 2000b : 12-13).

professionalism in the evaluation practice as possible.

(2) Evaluation Unit : In the KISTEP, the Programme Coordinator for Space Science and Fundamental Research has a general responsibility for the overall management of the CRI programme. In line with this role, the coordinator prepared the evaluation plan and managed the whole of the detailed evaluation process (KISTEP, 2002b). Therefore, it is evident that this coordinator acted as the evaluation unit in this evaluation, even though he received a great deal of assistance from the other PCs concerned. As a result, it is believed that the designation of the evaluation unit and its actual work were relatively rational and good, in that the coordinator had the expertise necessary for the overall management of this programme.

(3) Evaluation Information : The principal evaluation information was drawn from the materials submitted by the project leader. These included quantitative information, including articles published in SCI listed journals and invitations to international conferences, as well as qualitative information such as the research plan for the second stage, in accordance with the four evaluation criteria. In particular, in reflecting the initial objectives of the programme, much information from the high quality research articles was needed. Therefore, it appears that while qualitative and quantitative information were equally utilised, quantitative information was considered more important than in other Korean evaluation exercises mentioned above.

Consequently, it is clear that the entire evaluation resources were employed and invested at the proper level. In fact, the evaluation manpower has a very distinct characteristic. Unlike in other Korean evaluation exercises, many prominent international scientists were involved in the evaluation of the CRI programme, as a part of the government efforts to secure a high level of fairness and professionalism in the evaluation findings. In addition, regarding the evaluation information, much more quantitative information was required in order to reflect the quality of the core research outputs.

#### **6.3.5.3. Evaluation Performance**

(1) Evaluator : Even though four groups of external evaluation manpower were involved in this evaluation, the evaluation results of both the evaluation panel and international

evaluators were directly reflected in the final findings (MOST and KISTEP, 2000a : 5 ; MOST and KISTEP, 2000b : 4, 7). The evaluation panel was entirely made up of private sector experts from academia, industry and the research institutes. Three prominent international experts evaluated each research project. They came from 12 countries, including the USA, Japan, the UK and Switzerland. In fact, 51 evaluators among a total of 77 were American scientists (MOST and KISTEP, 2002b : 6). From this, we may see that this evaluation was entirely conducted by external evaluators, including international evaluators, and none of them came from the government ministries or agencies.

(2) Evaluation Method : The analysis of the materials submitted by the project leader and the bibliometric analysis of the research outputs were adopted as the key evaluation methods. Both the evaluation panels and international experts evaluated the excellence and the originality of the research outputs of each project, on the basis of the qualitative analysis of evaluation materials (MOST and KISTEP, 2000b : 4). For the bibliometric analysis, in particular, calculations were based on research articles published in Science Citation Index (SCI) listed journals (MOST and KISTEP, 2000a : 6).

As a result, it is evident that a greater emphasis was placed on the bibliometrics of research articles than in other evaluation exercises. It seems that this resulted from the principle of the CRI evaluation exercise, the intention of which was to judge the quality of research outputs using the most easily quantified evaluation criteria and indicators possible.

(3) Evaluation Criteria and Indicators : In the 2000 evaluation of the CRI, four evaluation criteria and five evaluation indicators were adopted, relatively fewer than in other cases. Each evaluation criterion was given a different weight as follows : the achievement of research objectives (20%), the originality of research outputs (40%), potentials as world-class researcher (20%) and appropriateness of future research plans (20%) (MOST and KISTEP, 2000a : 5 ; MOST and KISTEP, 2000b : 5). It is thought that the adoption of fewer evaluation criteria and indicators represented the MOST's policy of evaluating the quality of core research outputs, using simplified and more easily quantified evaluation criteria and indicators (MOST and KISTEP, 2000a : 1). In addition, of the four evaluation criteria, more weight was placed on the originality of the research outputs.

In sum, it is clear that the three components of the evaluation performance were adequately selected and effectively implemented. In particular, the placing of greater emphasis on the articles of the SCI listed journals with a high impact factor in the bibliometric analysis is considered a distinctive feature of the evaluation method. Meanwhile, the adoption of fewer evaluation criteria and the placing of more weight on the originality of the research outputs were listed as other features of the evaluation criteria.

#### ***6.3.5.4. Evaluation Utilisation***

(1) Evaluation Report : The evaluation report was distributed internally to the related members of staff in the MOST and the KISTEP as well as to evaluators involved. However, it was not made available to the public either electronically on the website or in hard copy form. Therefore, there is a need for the final evaluation report to be made open to, at least, the related S&T communities, even though the principal purpose of this evaluation was restricted to the budget adjustment of R&D projects. This is because the programme covers very creative basic research that is interesting to many scientists and researchers in related research fields.

(2) Utilisation Type : In the MOST evaluation, it is assumed that around 15 or 20 % of the lower ranked projects will be cut (MOST, 2002f ; MOST, 2002g). In line with the general policy, of the 27 R&D projects under the CRI programme scheme, five that had a grading of either “fair” or “unsatisfactory” were terminated, whereas the rest continued to get research funds (MOST and KISTEP, 2000b : 12-13). In addition, it is expected that these evaluation findings were utilised to improve the future structure of the programme. From this, it is evident, at the current stage, that more focus was placed on short-term and instrumental utilisation like the termination of the project, rather than on conceptual utilisation such as the investigation of the technological trends and developments.

Consequently, there are some requirements for improvement, bearing in mind that the initial intention of this evaluation exercise concentrated on the managerial matters at the project level. Therefore, firstly, the targets of the dissemination of the evaluation report need to be extended to the research communities including the learned societies concerned. Then, it is also desirable that the utilisation type of evaluation findings should

be diversified to reflect a more long-term perspective.

#### **6.3.5.5. *Synthesis***

From the results of the meta-evaluation conducted above, it seems, at the current stage, that the overall evaluation system was relatively well established and its actual exercise was effective, even though there is little need to improve the evaluation utilisation. In particular, it is believed that, in this system, the evaluation manpower and the evaluators have characteristics that are very different those of other evaluation systems. In Korea, it is the first time that many distinguished international scientists directly participated in the evaluation process. So, it is expected that this successful experience will affect other evaluation systems and their exercises in both the MOST and other central ministries responsible for national R&D programmes, including the MOE, the MOCIE, the MIC and the MOHW.

However, there are some improvements necessary in relation to the evaluation utilisation. As already noted, in particular, it is assumed that these improvements should be necessary for the related S&T communities. This is because the evaluation findings of the CRI programme are supposed to guide the future direction and the technological trends in related research fields.

#### **6.3.6. *Section Conclusion***

On the basis of the results of the evaluating the five Korean evaluation exercises, it is apparent that, in general, four of these systems, i.e. excluding the evaluation system for the CRI programme, are at early stage in their development. Therefore, it seems currently that there are many improvements that need to be made in the directive components of R&D evaluation systems, for example, in the evaluation paradigm and evaluation utilisation, even though the instrumental and methodological components, including the evaluation resources and the evaluation performance, are relatively sufficient and suitable. This is mainly because these evaluation systems were introduced in the late 1990s and early 2000.

In addition, the Korean R&D evaluation exercises, excluding the CRI case, have the distinctive feature that they are carried out every year. There is no doubt that such a short interval between evaluations has had a considerable influence on the overall evaluation system and its actual exercise. As a result, at the current stage, there are many improvements that need making in most Korean evaluation systems. In what follows, some characteristics and limitations of the Korean R&D evaluation system will be discussed by reference to some sub components of the meta-evaluation model.

Firstly, in four out of the five evaluation exercises, i.e. excluding the evaluation of the CRI programme, the evaluation purposes need to be expanded to include knowledge enhancement, even though this may be expressed in a somewhat implicit and indirect way. This is because these evaluation exercises normally deal with large-scale policies and programmes that have a great deal of long term influence on S&T and R&D areas as well as many government-supported research institutes that are considered as principal performers in the public sector R&D activities. For the same reason, an impact evaluation needs to be added to the four evaluation systems other than the evaluation of the CRI programme, as a further evaluation type.

Meanwhile, there are two essential preconditions that must be met, if these improvements in the evaluation purposes and types are to be realised. First of all, the evaluation object should be re-established not at the short-term and operating level but at the more long-term and strategic level. In particular, in the evaluation system of the major policy issues, there is a strong need for the evaluation objects to be upgraded from current sub programmes or main projects level to policy itself level and also for them to be extended to include cross-departmental policies under the responsibility of S&T related government ministries such as the MOE, the MOCIE, the MIC and the MOHW. In addition, in the case of the evaluation system of national R&D programmes, it is also strongly recommended that the evaluation objects be extended from the current sub programmes and projects level to the R&D programme itself. As already discussed, because of the lower level of the evaluation objects, the two evaluation systems have more features of annual monitoring than of *ex post* evaluation.

Next, the related government policies, in particular regarding the research councils and their research institutes, should be transferred from their current emphasis on managerial



matters to an emphasis on the systematic and strategic R&D governance structure in the public sector at national level. This is because the two evaluation systems dealt with the performance and the future research strategy of the principal R&D performers in the public sector in Korea.

Secondly, it seems, at the current stage, that the evaluation resources used were sufficient and suitable, bearing in mind that the evaluation paradigm of most Korean evaluation systems was established largely at the operating level rather than the strategic level. However, it is believed that, in the future, the overall evaluation resources should be amended in accordance with the improved evaluation paradigm discussed above.

Despite the improvements needed for the future, at the current stage, some distinctions can be drawn in relation to the evaluation resources. Firstly, as in general cases of R&D evaluation, two types of evaluation manpower were involved in the five Korean evaluation exercises : firstly the internal staff of the evaluation unit and secondly the external experts making up the evaluation committee or panel. The external manpower was entirely made up of the private sector experts, including academia, industry, the research institutes and the media. This is regarded as one of the Korean government efforts to secure the fairness and objectiveness of the evaluation exercise by making as many external experts as possible participate in the actual evaluation process.

Another distinction can be drawn in relation to the evaluation information. In all five evaluation exercises, in order to obtain the key evaluation information, the evaluators relied mainly on second hand sources such as performance report submitted by the institutions to be evaluated. Thus, it seems that the evaluator made few efforts to actively gather the first hand evaluation information, for example by conducting extensive consultation with related people or commissioning the necessary research in advance of the actual evaluation exercise.

There is also one distinction to be made in relation to the evaluation unit. As in general cases of R&D evaluation, in the five Korean evaluation exercises, an independent organisation was designated as the evaluation unit. However, it acted only as the secretariat to the evaluation committee or evaluation panel, unlike in the UK evaluation cases, where the members of the evaluation unit were directly involved in the actual

evaluation of some sub evaluation objects.

Thirdly, it is thought that the evaluation performance was relatively effectively implemented, because the current evaluation paradigm was more or less inevitably established in order to realise the initial government intentions to focus on the managerial matters in relation to the policies, programmes and research institutions at the operating level. Despite this general effectiveness of the evaluation performance, there is one common feature of the five evaluation exercises that needs to be mentioned. That is, all five evaluations adopted a grading system, consisting of four or five grades, to examine individual evaluation objects or sub objects according to the evaluation criteria and indications.

In particular, the adoption of the international evaluators in the evaluation of the CRI programme is considered as a very distinctive feature in the evaluation performance. As already noted, the MOST invited many prominent international scientists to the 2000 CRI evaluation in order to secure its fairness and professionalism. Therefore, it is expected that the MOST's successful experience in relation to the international evaluators will affect other evaluation systems of the MOST as well as of other S&T related central ministries, including the MOE, the MOCIE, the MIC and the MOHW.

Meanwhile, in the evaluation of major policy issues, the meta-evaluative approach was adopted as the main evaluation method. In order to evaluate the overall policy process in relation to the pertinent policy issues, the members of the PEC firstly reviewed the self-evaluation report conducted by individual ministries using the same evaluation process and criteria as the PEC, and then judged the rationality and effectiveness of the self-evaluation activities. This is thought to be one of the efforts that the Korean government has made with the intention of enhancing the capability of the evaluation activities at each ministry level.

However, there is a need for improvement in the evaluation criteria used, in particular, in the evaluation of the member research institutes under the KOI. A total of 22 evaluation criteria and 45 indicators were adopted to evaluate the 10 sub evaluation objects. These numbers need to be reduced appropriately, because too many evaluation criteria and indicators place on administrative burden on the evaluators as well as on related staff and

institutions.

Finally, it is clear that, in the five evaluation exercises, much more emphasis was placed on the instrumental utilisation of the evaluation findings from the short-term and managerial perspectives than on their conceptual utilisation from the long-term and strategic perspectives. Therefore, the evaluation findings of the five R&D evaluation exercises are expected to contribute more to the immediate improvement of the related R&D programmes and the management system of the research institutions at ministry level than to the future development of S&T and R&D policies at national level. It appears that this situation had a close connection with the evaluation paradigm, and in particular the evaluation purposes and objects, which were defined at a rather low level.

In conclusion, it is apparent that, among the five Korean R&D evaluation systems, the 2000 evaluation of the CRI programme is the best of the examples of evaluation exercises in Korea. Meanwhile, both the evaluation of the national R&D programmes and the evaluation of the member research institutes under the KOCI could be considered as relatively fair cases. On the other hand, it seems that there are a great many improvements required in both the evaluation of major policy issues of the MOST and the evaluation of the five research councils.

#### **6. 4. Comparison of the UK and the Korean R&D Evaluation Systems**

In this section, I shall conduct a comparison of the UK and the Korean national R&D evaluation systems. For this comparison, in principle, the findings of the meta-evaluation exercises conducted in the previous two sections are utilised. In addition, in order to provide a more comprehensive understanding of the two national evaluation systems, the hierarchy of the national R&D evaluation systems and the evaluation timing, namely the frequency of evaluations, are used as useful extra information for the comparison, even though neither is included from the meta-evaluation components.

By comparing the two national R&D evaluation systems, firstly, any possibility of applying aspects of the UK R&D evaluation system to the Korean case will be identified and then, the extent to which the UK system and its practice could be applied to the

Korean case will be discussed. Therefore, along with the findings of the meta- evaluation, the results of this comparison will be essential references for Chapter 7, which will deal with the suggestions for the applicability of elements of the UK national R&D evaluation system to the Korean system.

As already noted, the comparison between the UK and the Korean exercises is being conducted in order to obtain the basic materials to identify any applicability of the UK cases to the Korean cases. For this reason, further discussion during the comparative works will firstly be directed at whether there are any differences between the UK system and the Korean system and then at why these differences between them have arisen. Therefore, any discussions concerning which exercise is better or more suitable will be excluded from this section, even though this type of discussion was partially or indirectly undertaken as a part of the meta-evaluation.

#### ***6.4.1. Comparison of the Structure of National R&D Evaluation System***

As already showed in (Figure 4-2) and (Figure 5-2) respectively, the UK and the Korean national R&D evaluation systems normally include five sub evaluation systems, which are considered as essential and practical elements at the national level. In what follows, firstly the current types of evaluation exercise conducted regularly in each country will be compared and then a comparison will be made of the relationship among the five evaluation exercises, with a particular focus on the inputs of evaluation findings of the lower level evaluation systems to the process of the upper level ones. Finally, I shall conduct a comparison of the evaluation timing of the five evaluation exercises in each country, even though it was excluded from both case studies and the meta-evaluation in each country. This is because evaluation timing is usually regarded as one of the factors influencing other evaluation components, including the evaluation object, the evaluation methods, the evaluation utilisation and so on (Lee et al., 1994a ; OPM of Japan, 1997).

##### ***6.4.1.1. Current Types of Evaluation Exercises***

At the moment, there are five sub evaluation systems in both the UK and the Korean national R&D evaluation system. So, each national system includes sub evaluation

systems concerning S&T and R&D policy, research institutions (i.e. one each for research councils and research institutes), S&T or R&D programmes and R&D projects. The detailed types of the evaluation systems and the actual sub evaluation exercises recently conducted are shown in <Table 6-1>.

**<Table 6-1> Current Types of R&D Evaluation Exercises in the UK and Korea**

UK Case	Type of Evaluation	Korean Case
Cross-cutting Review of Science and Research (2000 ; 2002)	<b>Policy Evaluation</b>	Evaluation of Major Policy Issues of the MOST (2001)
Quinquennial Review of the Six Grant Awarding Research Councils (2001)	<b>Institution Evaluation (Research Councils)</b>	Evaluation of the Five Research Councils (2001)
Institute Exercise Assessment (IAE) by the BBSRC (2001)	<b>Institution Evaluation (Research Institutes)</b>	Evaluation of Member Research Institutes by the KOCI (2002)
Evaluation of Funding for UK Civil Space Activity by the DTI (2001)	<b>Programme Evaluation</b>	Evaluation of National R&D Programmes (2001)
Review of the Gastrointestinal Response to Food (GRF) Programme by the BBSRC (2000)	<b>Project Evaluation</b>	Evaluation of the Creative Research Initiatives (CRI) by the MOST (2000)

From this, it is evident that both countries' national R&D evaluation systems were well organised and that their actual exercises were also appropriately implemented. This is because it is normally argued that the national R&D evaluation system of a certain country needs to include four types of the sub evaluation system : policy evaluation, institution evaluation, programme evaluation and project evaluation (Cabinet Office, 1989 ; Georghiou, 1989 : 18 ; Hills and Dale, 1995 : 43-44).

Therefore, in comparing the sub evaluation systems and their actual practices in the UK and Korea, there is no difference between them, at least from the perspective of the institutionalisation of the evaluation systems, even though there might be more dissimilarity between them at the actual stage of their implementation.

#### **6.4.1.2. Relations among Evaluation Exercises**

In principle, it is assumed that there are direct or indirect relations among the diverse sub

evaluation systems making up a national R&D evaluation system. That is, these comprise the upper and lower parts in the overall R&D evaluation system as a whole. Therefore, at the national level, the evaluation findings of the lower parts, for example the evaluation of the research institutions concerned and the evaluation of S&T or R&D programmes, is expected to be one of the essential inputs to the evaluation of science and research policies. Meanwhile, it is desirable that, in each central government organisation, the results of the R&D project evaluations should be reflected in the process of the R&D programme evaluation and the related institution evaluations. In the light of this, the detailed relationships among individual sub evaluation exercises in the UK and Korea are summarised in <Table 6-2>. In this table, a solid line represents a direct relation, whereas a dotted line describes an indirect one.

<Table 6-2> Relations among Sub Evaluation Exercises in the UK and Korea

UK Case		Korean Case	
Type of Evaluation	Relationship	Relationship	Type of Evaluation
Cross-cutting Review of Science and Research	↑↑↑	↑↑	Evaluation of Major Policy Issues of the MOST
Quinquennial Review of the Six Grant Awarding Research Councils	↑↑	↑	Evaluation of Five Research Councils
Institute Assessment Exercise (IAE) by the BBSRC	↑	↑	Evaluation of Member Research Institute by the KOCI
Evaluation of Funding for UK Civil Space Activity by the DTI		↓↑	Evaluation of National R&D Programmes
Review of the Gastro-intestinal Response to Food (GRF) Programme by the BBSRC		↑	Evaluation of the Creative Research Initiatives (CRI) by the MOST

From this, it is evident that, in the UK evaluation system, the final findings of the four evaluation exercises making up lower parts in the national R&D evaluation system, namely, the quinquennial review of the six grant awarding research councils, the IAE conducted by the BBSRC, the evaluation of the funding for UK civil space activity and the review of the GRF programme by the BBSRC, were directly or indirectly reflected in the process of the cross-cutting review of science and research which is conducted at the highest level of the national R&D evaluation system. Therefore, it is apparent that there

are relatively close relationships among the sub evaluation exercises in the UK national R&D evaluation system.

On the other hand, it seems that, in the Korean system, the policy evaluation which was represented by the evaluation of the major policy issues of the MOST, was conducted separately from other evaluation exercises, even though the final findings of two of those exercises, namely, the evaluation of the national R&D programmes and the evaluation of the CRI programmes of the MOST, were part of the input to that process. In addition, another two evaluation systems that deal mainly with the R&D activities implemented by many GRIs, namely the evaluation of the national R&D programmes and the evaluation of the GRIs under the research councils, have been independently conducted from the perspective of each evaluation unit's own needs, even though, in a recent development, the former evaluation system has started to utilise some results of the latter one. Therefore, there has been some criticism of the redundancy and overlapping that exists between these two evaluation systems. Thus, it appears that there is a lack of close linkage between the policy evaluation by the OPM/the OGPC, the programme evaluation by the NSTC and the evaluation of research councils by the OGPC, which are considered as upper level evaluation exercises in the related areas.

Consequently, it is believed that the UK case is superior to the Korean one, in terms of the relationships between the sub evaluation exercises of the national R&D evaluation system.

#### ***6.4.1.3. Evaluation Timing***

In general, it is thought that the evaluation timing, namely the frequency of evaluation exercise, is one of the main factors influencing other evaluation components, including the evaluation purposes, the evaluation object and evaluation utilisation, although there has not been much previous discussion of this subject (Lee et al., 1994a ; OPM of Japan, 1997). In particular, it is clear that a shorter period between evaluations will have the effect that the evaluation purposes are determined from a short-term and instrumental perspective, whereas a longer period between them will have the opposite effect. Meanwhile, in general, it is anticipated that the intervals between policy evaluations will be longer than those between project evaluations. This is because the policy evaluation

normally deals with more complex and wide ranging affairs that are expected to have a considerable influence on S&T and R&D areas as a whole.

For this reason, I shall discuss the evaluation timing as one of the three elements in a comparison of the structure of the national R&D evaluation systems of the UK and Korea. The current frequency of the five evaluation exercises in the UK and Korea is shown in <Table 6-3>.

**<Table 6-3> Evaluation Timing of Sub Evaluation Exercises in the UK and Korea**

UK Case		Korean Case	
Type of Evaluation	Frequency	Frequency	Type of Evaluation
Cross-cutting Review of Science and Research	Biannual	Six Monthly	Evaluation of Major Policy Issues of the MOST
Quinquennial Review of the Six Grant Awarding Research Councils	Every Five Years	Every Year	Evaluation of the Five Research Councils
Institute Assessment Exercise (IAE) by the BBSRC	Every Four Years	Every Year	Evaluation of Member Research Institutes by the KOCI
Evaluation of Funding for UK Civil Space Activity by the DTI	Covering some 25 Years	Every Year	Evaluation of National R&D Programmes
Review of the Gastro-intestinal Response to Food (GRF) Programme by the BBSRC	On Completion (in principle)	3 Years after launch	Evaluation of the Creative Research Initiatives (CRI) By the MOST

From this, it is evident that most UK evaluation exercises are conducted at relatively longer intervals, for example, at least from every two years to a maximum of every five years. In fact, the evaluation exercises of the research councils and their research institutes have been conducted every four or five years. It is clear that a relative longer intervals between evaluations in the UK arise from the long history of its government having to make efforts to provide the SET communities with as much autonomy and stability as possible (MoR, 1918 ; Flanagan and Keenan, 1998 : 37).

On the other hand, four out of five Korean evaluation exercises, i.e. excluding the evaluation of the CRI programme, are conducted every year or on a six monthly basis. It seems that this situation has arisen partially due to the fact that, since these evaluation



systems were introduced in the late 1990s and early 2000, the Korean government wanted to settle them down as soon as possible. As a result, there has been a greater criticism that too frequent evaluation exercises have put considerable administrative burden on both the research institutions and the researchers to be evaluated and that, in turn, this situation has been one of the reasons for the lower reliability of the evaluation exercises themselves as well as for the weaker utilisation of evaluation findings (ECRC and KISTEP, 2001 : 121 ; KOCI, 2002c ; 81-83).

#### ***6.4.2. Comparison of the Evaluation System for S&T and R&D Policy***

As already noted, “the cross-cutting review of science and research in 2000” was selected as the case study of an evaluation exercise in relation to S&T and R&D policy in the UK, whereas, in Korea, “the evaluation of the major policy issues of the MOST in 2001” was selected. In what follows, the two countries’ evaluation systems of S&T and R&D policy will be compared on the basis of both the case studies in the previous chapters and the meta-evaluations in the previous sections. Through this comparison, I intend to obtain useful information for the following chapter which will deal with suggestions for applying elements of the UK case to the Korean situation.

##### ***6.4.2.1. Evaluation Paradigm***

The comparison of the three evaluation components making up the evaluation paradigm of S&T and R&D policy in the UK and Korea is summarised in <Table 6-4>.

From this, it is clear that there is no difference in the evaluation purposes in that the two cases emphasised both the accountability of related government ministries and the improvement of the policy and programmes concerned. Unlike in the UK case, in the Korean case, the interim evaluation was adopted as a main evaluation type so that there is a slight difference in relation to this component.

A considerable difference can be identified between the evaluation objects in the two cases. The UK case involved both the SET policies and the main programmes implemented by the central government departments concerned, including the DTI/the

OST, the DfES, the DEFRA and the DfT, as well as by their main NDPBs, such as the research councils and the funding councils. On the other hand, the Korean case included the sub programmes and their main projects comprising the pertinent policy conducted by only one central ministry, the MOST. Therefore, in fact, there are some limitations in that the Korean case cannot evaluate the related policy itself as a whole and that it also cannot manage to handle the S&T and R&D policies from the broader perspective at national level.

**<Table 6-4> Comparison of Evaluation Paradigm for S&T and R&D Policy in the UK and Korea**

UK Case	Components	Korean Case
- accountability - improvement of policy/programme	<b>Evaluation Purpose</b>	- accountability - policy improvement
- SET policies - main S&T and R&D programmes <i>* covered policies and programmes of both SET related departments and their main NDPBs</i>	<b>Evaluation Object</b>	- sub programmes comprising the pertinent policy issues - main projects <i>* included only the policies of one ministry (i.e. MOST)</i>
- output evaluation - impact evaluation (where appropriate)	<b>Evaluation Type</b>	- interim evaluation - output evaluation - impact evaluation (where appropriate)

Consequently, in comparing the evaluation paradigm of the S&T and R&D policy in the UK and Korea, it is apparent that, in particular, the UK evaluation objects were better established than those in Korea, although there is no considerable difference between the two other evaluation components, namely the evaluation purpose and the evaluation types.

#### **6.4.2.2. Evaluation Resources**

A comparison of the three evaluation components making up the evaluation resources of S&T and R&D policy in the UK and Korea is shown in <Table 6-5>.

From this, it appears that there are some differences in relation to evaluation manpower as well as evaluation information, whereas there is a similarity between the evaluation units of the two systems. One distinction can be observed in relation to evaluation manpower,

in particular, external manpower. While the UK external manpower came from the private sector, including the universities and industry, as well as from the government departments relevant to SET affairs such as the DTI and the DfES, the Korean external manpower, namely members of the PEC and the support team, came entirely from the private sector, including academia, industry, the research institutes and the media. It is believed that the particular characteristics of the UK external evaluation manpower are influenced by the UK government policy that “the customer-contractor principle” should operate in applied research and development carried out or sponsored by the government, whether by the research councils or elsewhere, which had been recommended by the Rothschild Report in 1971 (Rothschild, 1971).

**<Table 6-5> Comparison of Evaluation Resources for S&T and R&D Policy in the UK and Korea**

UK Case	Components	Korean Case
<ul style="list-style-type: none"> <li>- internal manpower <ul style="list-style-type: none"> <li>• staff of the Enterprise Team in HM Treasury (1)</li> </ul> </li> <li>- external manpower (5) <ul style="list-style-type: none"> <li>• from DTI, DfES, universities, and private sector</li> </ul> </li> </ul>	<b>Evaluation Manpower</b>	<ul style="list-style-type: none"> <li>- internal manpower <ul style="list-style-type: none"> <li>• staff of the AMO for PAE under the OGPC (35)</li> </ul> </li> <li>- external manpower <ul style="list-style-type: none"> <li>• members of the PEC (30)</li> <li>• support team (29)</li> </ul> </li> </ul>
<ul style="list-style-type: none"> <li>- the Enterprise Team in HM Treasury</li> <li><i>* one of the 19 spending teams</i></li> </ul>	<b>Evaluation Unit</b>	<ul style="list-style-type: none"> <li>- the AMO for PAE of the OGPC</li> <li>• Director-General I for PAE</li> <li><i>* responsible for the economic policies</i></li> </ul>
<ul style="list-style-type: none"> <li>- existing research available</li> <li>- extensive consultation</li> </ul>	<b>Evaluation Information</b>	<ul style="list-style-type: none"> <li>- performance report</li> <li>- self-evaluation report</li> <li>- extensive consultation (supplementary sources)</li> </ul>

Meanwhile, there is another difference concerning evaluation information. The UK system utilised relatively objective information that had been produced by outside researchers, mainly from the universities, or created by the evaluators themselves, whereas the Korean exercise relied mainly on somewhat subjective information that had been submitted by the pertinent ministries to be evaluated. It seems that this distinctive feature of the Korean case was closely connected with the adoption of the meta-evaluative approach as an evaluation methods, as discussed in the following sub-section on evaluation performance.

### 6.4.2.3. Evaluation Performance

The comparison of the three evaluation components, consisting of the evaluation performance in relation to S&T and R&D policy in the UK and Korea, is summarised in <Table 6-6>.

**<Table 6-6> Comparison of Evaluation Performance for S&T and R&D Policy in the UK and Korea**

UK Case	Components	Korean Case
<ul style="list-style-type: none"> <li>- steering group <ul style="list-style-type: none"> <li>• staff from HM Treasury, DTI, DfES</li> <li>• advisers from academia and industry</li> </ul> </li> </ul>	<b>Evaluator</b>	<ul style="list-style-type: none"> <li>- members of the PEC <ul style="list-style-type: none"> <li>• entirely made up of private sector people</li> <li>• professional assistance from the support team</li> </ul> </li> </ul>
<ul style="list-style-type: none"> <li>- literature review</li> <li>- extensive consultation</li> </ul>	<b>Evaluation Method</b>	<ul style="list-style-type: none"> <li>- review of performance reports &amp; self-evaluation reports <ul style="list-style-type: none"> <li>* <i>meta-evaluative approach adopted</i></li> </ul> </li> <li>- site visits &amp; consultation (supplementary method) <ul style="list-style-type: none"> <li>* <i>grading system adopted</i></li> </ul> </li> </ul>
<ul style="list-style-type: none"> <li>- not clear about evaluation criteria</li> <li>- four types of evaluation indicators <ul style="list-style-type: none"> <li>• mixed qualitative and quantitative ones</li> </ul> </li> </ul>	<b>Evaluation Criteria &amp; Indicators</b>	<ul style="list-style-type: none"> <li>- 6 common evaluation criteria <ul style="list-style-type: none"> <li>• mainly qualitative ones</li> </ul> </li> <li>- evaluation indicator <ul style="list-style-type: none"> <li>• not clear</li> </ul> </li> </ul>

Some features of the two evaluation systems can be drawn from this summary. First of all, there is one difference between two systems regarding the evaluators. Even though the two evaluation exercises were conducted only by the external evaluators, the detailed member of the evaluators was different. As already indicated, while evaluators from both the private sectors and the irrelevant government departments conducted the UK evaluation exercise, evaluators only from the private sectors conducted the Korean exercise.

One distinctive feature of the evaluation method in the Korean exercise was that a meta-evaluative approach was adopted. That is, the Korean exercise firstly examined the self-evaluation reports submitted for evaluation in advance by individual ministries and then judged the rationality and the effectiveness of the overall system and the actual

implementation of the self-evaluation activities (OGPC and PEC, 2001a ; PEC, 2002b). This is considered as one of the efforts made by the Korean government to improve the capability of evaluation activities at ministry level.

On the other hand, no difference was identified between the UK and the Korean exercises in relation to the evaluation criteria and indicators.

#### **6.4.2.4. Evaluation Utilisation**

The comparison of the two evaluation components making up the evaluation utilisation of S&T and R&D policy in the UK and Korea is summarised in <Table 6-7>.

**<Table 6-7> Comparison of Evaluation Utilisation for S&T and R&D Policy in the UK and Korea**

<b>UK Case</b>	<b>Components</b>	<b>Korean Case</b>
<ul style="list-style-type: none"> <li>- published</li> <li>- available on the website</li> </ul>	<b>Evaluation Report</b>	<ul style="list-style-type: none"> <li>- published <ul style="list-style-type: none"> <li>• a comprehensive report</li> <li>• individual ministries' reports</li> </ul> </li> <li>- available on the website</li> </ul>
<ul style="list-style-type: none"> <li>- instrumental utilisation <ul style="list-style-type: none"> <li>• reflected in 2000 SR settlement</li> </ul> </li> <li>- conceptual utilisation <ul style="list-style-type: none"> <li>• inputs to the 2000 Science and Innovation White Paper</li> </ul> </li> </ul>	<b>Utilisation Type</b>	<ul style="list-style-type: none"> <li>- instrumental utilisation <ul style="list-style-type: none"> <li>• giving awards</li> <li>• budget adjustments</li> <li>• suggesting corrective measures and audit</li> </ul> </li> <li>- conceptual utilisation <ul style="list-style-type: none"> <li>• not considered</li> </ul> </li> </ul>

There is a great similarity between the UK and the Korean case in relation to the evaluation report in that both evaluation exercises published their final evaluation reports and also posted them on the website of the related organisations. However, there is a slight difference in the utilisation type of the evaluation findings. While both instrumental and conceptual utilisations were adopted with nearly equal weight in the UK system, greater emphasis was placed on instrumental utilisation in the Korean system

### 6.4.3. Comparison of the Evaluation System for S&T or R&D Programme

While “the evaluation of funding for UK civil space activities by the DTI in 2001” was selected as the case study of the evaluation exercise of S&T or R&D programme in the UK, “the evaluation of national R&D programmes in 2001” was chosen in Korea. In what follows, a similar discussion structure as that used in relation to the evaluation of S&T and R&D policy will be adopted in order to compare the evaluation exercises regarding S&T or R&D programmes in the UK and Korea.

#### 6.4.3.1. Evaluation Paradigm

The comparison of the three evaluation components making up the evaluation paradigm of S&T or R&D programmes in the UK and Korea is summarised in <Table 6-8>.

**<Table 6-8> Comparison of Evaluation Paradigm for S&T or R&D Programme in the UK and Korea**

UK Case	Components	Korean Case
- accountability of past expenditure - programme improvement	<b>Evaluation Purpose</b>	- accountability - programme improvement - knowledge sharing for R&D management in public sector
- programme-level works - policy-level work (where appropriate) * - <i>home programmes</i> - <i>international programmes</i>	<b>Evaluation Object</b>	- sub R&D programmes (161) - sub R&D projects (9,533)
- output evaluation - impact evaluation - assessment of future programme improvement	<b>Evaluation Type</b>	- interim evaluation - output evaluation

The most distinctive difference between the UK and Korean systems can be observed in relation to the evaluation objects. The evaluation objects of the UK exercise were set at the level of programmes and, where appropriate, a certain policy, and also covered both home and international programmes. As a result, the evaluation objects selected in the UK system adequately reflected the fact that the civil space activities need to be carried out as inter-departmental missions and as collaborative works with international partners.

On the other hand, in the Korean exercise, the evaluation objects were determined at the level of sub R&D programmes and projects.

It is clear that the difference in the evaluation objects actually selected must have affected the adoption of the evaluation types. While the UK exercise rationally adopted the three evaluation types necessary for the evaluation of large-scale and long-term S&T programme, the Korean exercise placed more emphasis on the interim and the output evaluation than on other types of evaluation such as impact evaluation and assessment of future strategies.

Meanwhile, there is no notable difference between the evaluation purposes of the two exercises, even though the Korean system proposed a knowledge sharing for R&D management in the public sector as an extra evaluation purpose.

#### ***6.4.3.2. Evaluation Resources***

A comparison of the three evaluation components comprising the evaluation resources of S&T or R&D programme in the UK and Korea is summarised in <Table 6-9>.

From this, it is evident that there is one difference between the UK and the Korean case in relation to the mission of the evaluation units. In the UK system, the TESE Directorate, one of many executive directorates in the DTI, made a master plan for the evaluation, managed the overall evaluation process and also conducted the actual evaluation. However, in the Korean system, the two evaluation units varied in their roles : the Science and Technology Assessment Division of the MOST was responsible for preparing the master plan for the evaluation work, while the Division of National R&D Evaluation of the KISTEP was mainly responsible for implementing the master plan at operating level. In addition, neither of them was involved in the actual conduct of the evaluation. As a result, the Korean evaluation units acted only as organiser of the evaluation or as the secretariat to the evaluation committee.

Another difference can be identified in the sources of evaluation information. The UK exercise utilised various information sources obtained by outside consultants as well as by the evaluators themselves, whereas the Korean exercise relied mainly on more or less

simple and, sometimes, subjective information sources that had been produced by the pertinent ministries and programme managers that were to be evaluated.

**<Table 6-9> Comparison of Evaluation Resources for S&T or R&D Programme in the UK and Korea**

UK Case	Components	Korean Case
<ul style="list-style-type: none"> <li>- internal manpower <ul style="list-style-type: none"> <li>• staff of TESE in DTI (3)</li> </ul> </li> <li>- external manpower <ul style="list-style-type: none"> <li>• specialists from Technopolis (5)</li> <li>• international space experts (3)</li> </ul> </li> </ul> <p><i>* contract basis</i></p>	<b>Evaluation Manpower</b>	<ul style="list-style-type: none"> <li>- internal manpower <ul style="list-style-type: none"> <li>• staff of MOST (8) and KISTEP (21)</li> </ul> </li> <li>- external manpower <ul style="list-style-type: none"> <li>• evaluation committee <ul style="list-style-type: none"> <li>→ steering committee (19)</li> <li>→ programme evaluation committees (83)</li> <li>→ project review committees (140)</li> </ul> </li> <li>• support team (16)</li> </ul> </li> </ul>
<ul style="list-style-type: none"> <li>- TESE Directorate in DTI <ul style="list-style-type: none"> <li>• Assessment Unit (AU)</li> </ul> </li> </ul>	<b>Evaluation Unit</b>	<ul style="list-style-type: none"> <li>- S&amp;T Assessment Division in MOST</li> <li>- Division of National R&amp;D Evaluation in KISTEP</li> </ul>
<ul style="list-style-type: none"> <li>- newly commissioned study</li> <li>- meetings with related people</li> <li>- extensive structured interviews</li> <li>- review of archive files</li> <li>- review of other ESA cases</li> </ul>	<b>Evaluation Information</b>	<ul style="list-style-type: none"> <li>- performance report submitted by ministry</li> <li>- presentation by programme manager</li> </ul>

Unlike in the case of the evaluation unit and information, there is no difference between the two systems in relation to evaluation manpower in that both internal and external evaluation manpower was involved in the two evaluation processes.

#### **6.4.3.3. Evaluation Performance**

A comparison of the three evaluation components making up the evaluation performance of S&T or R&D programme in the UK and Korea is summarised in <Table 6-10>.

First of all, there is one striking difference regarding the evaluators. In the UK, it was mainly internal evaluators who led the evaluation exercise, although they received useful professional assistance from the outside specialists of both Technopolis and Scientific



Generics, which were involved in the evaluation on the basis of a temporary contract. On the other hand, the Korean exercise was entirely conducted by the members of the various evaluation committees, all of whom came from the private sector, namely from academia, industry, the research institutes and the media.

**<Table 6-10> Comparison of Evaluation Performance for S&T or R&D Programme in the UK and Korea**

UK Case	Components	Korean Case
<ul style="list-style-type: none"> <li>- internal evaluator <ul style="list-style-type: none"> <li>• three staff in AU under TESE</li> </ul> </li> </ul>	<b>Evaluator</b>	<ul style="list-style-type: none"> <li>- evaluation committee <ul style="list-style-type: none"> <li>• 8 programme evaluation committees</li> <li>• 14 project review committees</li> </ul> </li> </ul>
<ul style="list-style-type: none"> <li>- field survey</li> <li>- literature review</li> <li>- in-depth interview</li> <li>- comparative study between the UK and other ESA countries</li> </ul>	<b>Evaluation Method</b>	<ul style="list-style-type: none"> <li>- review of performance report</li> <li><i>* a grading system adopted</i></li> </ul>
<ul style="list-style-type: none"> <li>- six evaluation criteria <ul style="list-style-type: none"> <li>• rationale of public intervention</li> <li>• additionality of achievements</li> <li>• existence of alternatives</li> <li>• principal beneficiaries</li> <li>• appropriateness of funding among programmes</li> <li>• balance between home and international programmes</li> </ul> </li> </ul>	<b>Evaluation Criteria</b>	<ul style="list-style-type: none"> <li>- programme evaluation <ul style="list-style-type: none"> <li>• appropriateness of objectives</li> <li>• efficiency of implementation</li> <li>• level of performance/outputs</li> </ul> </li> <li>- project review <ul style="list-style-type: none"> <li>• redundancy of projects</li> <li>• needs for integration/linkage among projects</li> </ul> </li> </ul>

There is also a slight difference between the two countries' system concerning the evaluation method. In the UK exercise, diverse types of evaluation methods were adopted, and these were closely linked with the evaluation information available. However, the Korean exercise mainly adopted a literature review by reviewing the performance report submitted by each ministry. As a result, it is evident that the two evaluation systems adopted a qualitative approach as the main evaluation method, even though the details of the approach type varied depending on the evaluation information available.

Regarding the evaluation criteria, there is a great deal of similarity in that the two exercises mainly adopted qualitative evaluation criteria, even though their detailed contents were different.

#### 6.4.3.4. Evaluation Utilisation

A comparison of the two evaluation components of the evaluation utilisation of S&T or R&D programme in the UK and Korea is summarised in <Table 6-11>.

**<Table 6-11> Comparison of Evaluation Utilisation for S&T or R&D Programme in the UK and Korea**

UK Case	Components	Korean Case
- published • CD-ROM and hard copy - available on the website	<b>Evaluation Report</b>	- published - available on the website
- instrumental utilisation • improvements of space R&D programmes by public funding - conceptual utilisation • long-term development of space related UK industries <i>* most recommendations accepted by related departments</i>	<b>Utilisation Type</b>	- instrumental utilisation • budget adjustment • recommendations for improvement of programmes and projects - conceptual utilisation • not considered

There is a great deal of similarity between the two countries in the production and the dissemination of evaluation reports, because both evaluation exercises published their final reports and also posted them on the website of the related organisations. However, a slight distinction can be found in the utilisation type of the evaluation findings. While the evaluation findings were equally utilised in both the instrumental and conceptual perspective in the UK system, a greater emphasis was placed on instrumental utilisation in the Korean system.

#### 6.4.4. Comparison of the Evaluation System for Research Councils

“The quinquennial review of the six grant awarding research councils in 2001” was selected as the case study of the evaluation exercise of research councils in the UK, whereas, in Korea, “the evaluation of the five research councils in 2001” was chosen. In what follows, a similar discussion structure to that for other evaluation exercises will be adopted in order to compare the evaluation exercises of the research councils in the UK and Korea.

#### 6.4.4.1. Evaluation Paradigm

The comparison of the three evaluation components making up the evaluation paradigm of the research councils in the UK and Korea is summarised in <Table 6-12>.

**<Table 6-12> Comparison of Evaluation Paradigm for Research Councils in the UK and Korea**

UK Case	Components	Korean Case
<ul style="list-style-type: none"> <li>- accountability</li> <li>- improvement of research councils' management</li> <li>- knowledge enhancement (explicitly announced)</li> </ul>	<b>Evaluation Purpose</b>	<ul style="list-style-type: none"> <li>- accountability</li> <li>- improvement of research councils' management</li> </ul>
<ul style="list-style-type: none"> <li>- past performance (stage 1)</li> <li>- future management plan (stage 2)               <ul style="list-style-type: none"> <li>• mission/structure/governance</li> <li>• relations between RCs and their stakeholders</li> <li>• methods of priority setting and decision making</li> <li>• RCs' management and internal process</li> </ul> </li> </ul>	<b>Evaluation Object</b>	<ul style="list-style-type: none"> <li>- own activities and mission of supporting member GRIs (70%)</li> <li>- internal management system (30%)</li> </ul> <p><i>*10 sub evaluation objects → focused on managerial matters at the operating level</i></p>
<ul style="list-style-type: none"> <li>- output evaluation</li> <li>- assessment of future management</li> </ul>	<b>Evaluation Type</b>	<ul style="list-style-type: none"> <li>- output evaluation</li> <li>- assessment of R&amp;D plans &amp; GRI's development plan</li> </ul>

From this comparison, in principle, it is apparent that there are no differences between the two cases in relation to either the evaluation purpose and the evaluation type. In the evaluation systems in both the UK and the Korean cases, it is judged that the evaluation purposes were well established and that the evaluation types were also appropriately proposed in order to realise the initial policy objectives of conducting institution evaluation.

However, a considerable difference can be drawn between the evaluation objects of each country. As already discussed, it is clear that this distinction resulted mainly from the evaluation timing in that the UK exercise is in principle conducted every five years, whereas the Korean exercise is carried out every year. In the UK system, the evaluation objects were set up to deal with both the overall performance of the previous five years and the future management strategies and plans for the following five years. Also, the

individual main objects included affairs ranging from fundamental matters such as the mission, structure and governance of the research council system itself to managerial matters, including the individual research councils' internal management system. As a result, the evaluation objects in the UK exercise were established with a broader and more strategic perspective.

On the other hand, in the Korean system, the evaluation objects involved much past performance of the internal and external management of the research councils for only the previous one year, although these were intended to cover the future plans for their member research institutes' development. Therefore, it appears that the evaluation objects of the Korean exercise were more related to managerial matters at operating level, rather than to strategic and long-term affairs seen from the viewpoint of public R&D governance at the national level.

#### 6.4.4.2. Evaluation Resources

The results of the comparison of the three evaluation components making up the evaluation resources of the research councils in the UK and Korea, are shown in <Table 6-13>.

**<Table 6-13> Comparison of Evaluation Resources for Research Councils in the UK and Korea**

UK Case	Components	Korean Case
<ul style="list-style-type: none"> <li>- steering group (17)</li> <li>- review team (6)</li> <li>- 4 working groups (85 ; stage two)</li> <li><i>* mixed manpower with both internal and external experts</i></li> </ul>	<b>Evaluation Manpower</b>	<ul style="list-style-type: none"> <li>- internal manpower <ul style="list-style-type: none"> <li>• staff of DG for R&amp;D Support under OGPC (6)</li> </ul> </li> <li>- external manpower <ul style="list-style-type: none"> <li>• evaluation committee (10)</li> <li>• secretariat : KISTEP staff (2)</li> </ul> </li> <li><i>* contract basis</i></li> </ul>
<ul style="list-style-type: none"> <li>- Temporary mini Directorate under SEBG in the OST</li> </ul>	<b>Evaluation Unit</b>	<ul style="list-style-type: none"> <li>- Director-General for R&amp;D Support under OGPC <ul style="list-style-type: none"> <li>• 2 teams of all 4 teams</li> </ul> </li> </ul>
<ul style="list-style-type: none"> <li>- extensive public consultation</li> <li>- related literature and existing studies</li> <li>- in-depth discussion</li> </ul>	<b>Evaluation Information</b>	<ul style="list-style-type: none"> <li>- performance report</li> <li>- questionnaire survey (only reference information)</li> </ul>

From this comparison, it is evident that there are some differences in relation to both the evaluation manpower and evaluation information, even though there is no such difference in relation to the evaluation units. First of all, in the Korean exercise, the roles of the internal and external manpower were completely different. The internal manpower from the Director-General for R&D Support under the OGPC acted as the organiser of this evaluation work, whereas the actual responsibility for this evaluation was given only to the external manpower, namely the evaluation committee, who had secretarial assistance from the staff of the Division of the National R&D Evaluation in the KISTEP.

On the other hand, in the UK exercise, both internal manpower from the DTI, the OST and the research councils, and external manpower from other government departments, academia, research institutes, research charities and business were involved in the overall process of the evaluation. Consequently, several members of staff from the DTI, the OST, the six research councils and some research institutes that had a direct or indirect interest in the quinquennial review participated in the evaluation process, through active involvement in the steering group, the review team and 4 working groups. This is because the UK government has a policy to reflect as many diverse opinions of main stakeholders as possible in the evaluation process, even though they may belong to the organisations to be evaluated.

Another distinction between the two cases can be identified in the evaluation information used. The UK exercise utilised not only second hand information from the existing literature and studies but also more first hand information from extensive public consultations and in-depth interviews conducted by the evaluators themselves. Meanwhile, the Korean exercise relied mainly on relative second hand information source submitted by the research councils to be evaluated.

#### ***6.4.4.3. Evaluation Performance***

A comparison of the three evaluation components making up the evaluation performance of research councils in the UK and Korea is summarised in <Table 6-14>.

In the comparison between the two countries of the three components of evaluation performance, there is firstly a distinction in relation to the evaluators. As already

indicated, the UK exercise was conducted by both internal and external evaluators, whereas the Korean exercise was carried out only by external evaluators.

**<Table 6-14> Comparison of Evaluation Performance for Research Councils in the UK and Korea**

UK Case	Components	Korean Case
<ul style="list-style-type: none"> <li>- review team <ul style="list-style-type: none"> <li>• from each of DTI, OST, RC, research charity, business</li> </ul> </li> <li>- working group (only stage two) <ul style="list-style-type: none"> <li>• mainly from academia, RCs governments, research institutes, business</li> </ul> </li> </ul>	<b>Evaluator</b>	<ul style="list-style-type: none"> <li>- members of evaluation committee <ul style="list-style-type: none"> <li>• from academia, industry, research institute and the media</li> </ul> </li> </ul>
<ul style="list-style-type: none"> <li>- extensive public consultation <ul style="list-style-type: none"> <li>• including public consultation in each stage</li> </ul> </li> <li>- literature review</li> <li>- in-depth discussion</li> </ul>	<b>Evaluation Method</b>	<ul style="list-style-type: none"> <li>- review of performance report</li> <li>- site visit</li> <li>- conduct of questionnaire survey</li> <li><i>* grading system adopted</i></li> </ul>
<ul style="list-style-type: none"> <li>- main evaluation criteria <ul style="list-style-type: none"> <li>• appropriateness of organisational status</li> <li>• efficiency of management</li> <li>• effectiveness of management</li> </ul> </li> <li>- not clear about indicator</li> </ul>	<b>Evaluation Criteria &amp; Indicators</b>	<ul style="list-style-type: none"> <li>- 21 evaluation criteria</li> <li>- 36 evaluation indicators</li> <li><i>* qualitative criteria and indicators mainly adopted</i></li> </ul>

Regarding the evaluation method, the UK exercise had one distinctive feature. In the quinquennial review of the NDPBs and agencies of the UK, a public consultation has to be conducted, according the Cabinet Office guidelines (Cabinet Office, 2000a). Thus, in each stage of this review, both the review team and the working group undertook a public consultation exercise. Through this work, the evaluators, namely the review team and the working group, were able to identify various opinions of the key stakeholders, including the government departments concerned, the universities, the PSREs, the research charities and industry.

Meanwhile, there is a similarity between the two cases in relation to the evaluation criteria and indicators, in that qualitative ones were mainly adopted in both, even though they were different in detail.

#### 6.4.4.4. *Evaluation Utilisation*

The results of the comparison of the two components making up the evaluation utilisation of the research councils in the UK and Korea are shown in <Table 6-15>.

**<Table 6-15> Comparison of Evaluation Utilisation for Research Councils in the UK and Korea**

UK Case	Components	Korean Case
- published - available on the website	<b>Evaluation Report</b>	- published - not available on the website
- instrumental utilisation • input to 2002 cross-cutting review of science and research - conceptual utilisation • policy means for re-arranging roles and missions of RCs	<b>Utilisation Type</b>	- instrumental utilisation • notified to MPB and NSTC • budget adjustments - conceptual utilisation • not considered

As in other evaluation cases, a slight difference can be found in the utilisation type of the evaluation findings. While, in the UK system, the evaluation findings were equally utilised for both instrumental and conceptual purposes, in the Korean system, instrumental utilisation was considered as the principal utilisation type.

#### 6.4.5. *Comparison of the Evaluation System for Research Institutes under the Research Councils*

“The institute assessment exercise (IAE) by the BBSRC in 2001” was selected as the case study of the evaluation exercise of research institutes under the responsibility of the research councils in the UK, whereas “the evaluation of the member research institutes by the KOCI in 2002” was chosen in Korea. As a part of this work, nine research institutes were evaluated in the UK and seven research institutes were investigated in Korea. In what follows, a similar discussion structure to that for other evaluation exercises will be used in order to compare the evaluation exercises of the research institutes in the UK and Korea.

### 6.4.5.1. Evaluation Paradigm

A comparison of the three evaluation components making up the evaluation paradigm of research institutes in the UK and Korea is summarised in <Table 6-16>.

**<Table 6-16> Comparison of Evaluation Paradigm for Research Institutes in the UK and Korea**

UK Case	Components	Korean Case
<ul style="list-style-type: none"> <li>- accountability</li> <li>- improvement of the expenditure of research funds</li> </ul>	<b>Evaluation Purpose</b>	<ul style="list-style-type: none"> <li>- improvement of management system &amp; research programme</li> <li>- accountability</li> <li>- knowledge dissemination</li> </ul>
<ul style="list-style-type: none"> <li>- quality of science (QS)</li> <li>- knowledge transfer (KT)</li> <li>- competitive funding (CF)</li> <li>- visiting group (VG) assessment of overall activities of institute</li> <li>*- <i>past performance (QS, KT, CF, VG)</i></li> <li>- <i>future plan (KT, VG)</i></li> </ul>	<b>Evaluation Object</b>	<ul style="list-style-type: none"> <li>- R&amp;D matters (50%)</li> <li>- managerial matters (50%)</li> <li>* <i>10 sub evaluation objects</i></li> </ul>
<ul style="list-style-type: none"> <li>- output evaluation</li> <li>- assessment of research plan</li> </ul>	<b>Evaluation Type</b>	<ul style="list-style-type: none"> <li>- output evaluation</li> <li>- assessment of research plan</li> </ul>

From this, it is apparent that there is in principle no difference in relation to the evaluation purposes, even though these were slightly different in their detail. Also, the two exercises were very similar in adopting both output evaluation and the assessment of the research plan as evaluation types.

However, there were some difference in relation to the detailed contents and attributes of the evaluation objects, although these in principle included both research matters and internal managerial matters in the two evaluation systems. The UK exercise proposed four specific evaluation objects : the quality of science (QS), technology transfer (KT), competitive funding (CF) and the assessment of the overall activities of the institute by the visiting groups (VG), so that it was able to place greater emphasis on core research outputs and R&D activities themselves than on the R&D management system. Therefore, it appears that the UK government aimed to examine both the quality and level of key research performance supported by public funds as well as to improve the effectiveness of R&D expenditures by regularly conducting institution evaluations, namely the IAE.



On the other hand, the Korean exercise established two main evaluation objects in more or less broad terms, even though these were divided into 10 detailed sub evaluation objects. In addition, there were many more sub evaluation objects in relation to R&D management than in relation to research activities. As a result, it is apparent that the Korean exercise focused more on R&D management matters than on the research activities themselves and their quality and level.

#### 6.4.5.2. Evaluation Resources

The results of the comparison of the three components, making up the evaluation resources of the research institutes in the UK and Korea, are shown in <Table 6-17>.

**<Table 6-17> Comparison of Evaluation Resources for Research Institutes in the UK and Korea**

UK Case	Components	Korean Case
<ul style="list-style-type: none"> <li>- internal manpower <ul style="list-style-type: none"> <li>• staff of APU and Finance Group (FG) (6)</li> </ul> </li> <li>- external manpower <ul style="list-style-type: none"> <li>• 6 expert panels for QS (72) <ul style="list-style-type: none"> <li>→ from academia, business and government</li> </ul> </li> <li>• single panel for KT (12) <ul style="list-style-type: none"> <li>→ mainly from industry</li> </ul> </li> <li>• 9 visiting groups for VG (108) <ul style="list-style-type: none"> <li>→ from academia, business, government and overseas</li> </ul> </li> </ul> </li> </ul>	<b>Evaluation Manpower</b>	<ul style="list-style-type: none"> <li>- internal manpower <ul style="list-style-type: none"> <li>• staff of Evaluation and Management Team (3)</li> </ul> </li> <li>- external manpower <ul style="list-style-type: none"> <li>• evaluation committee (17) <ul style="list-style-type: none"> <li>→ 2 sub committees</li> </ul> </li> <li>• support team : STEPI staff (5) <ul style="list-style-type: none"> <li>* contract basis</li> </ul> </li> </ul> </li> </ul>
<ul style="list-style-type: none"> <li>- Assessment &amp; Policy Unit (APU) <ul style="list-style-type: none"> <li>• secretariat in 2001 IAE</li> <li>• evaluation of elements of QC</li> </ul> </li> </ul>	<b>Evaluation Unit</b>	<ul style="list-style-type: none"> <li>- Evaluation &amp; Management Team</li> </ul>
<ul style="list-style-type: none"> <li>- performance report <ul style="list-style-type: none"> <li>• very quantitative information</li> </ul> </li> <li>- site visit</li> </ul>	<b>Evaluation Information</b>	<ul style="list-style-type: none"> <li>- performance report <ul style="list-style-type: none"> <li>• mixed qualitative and quantitative information</li> </ul> </li> <li>- site visit</li> </ul>

From this, it is evident that there are some differences between the two cases in relation to evaluation manpower and the evaluation unit. First of all, in the Korean exercise, the roles of the internal and external manpower were completely different. The internal manpower from the Evaluation and Management Team (EMT) in the KOCI only organised the

evaluation work, whereas the actual responsibility for the evaluation was entirely placed on the external manpower, namely the evaluation committee, who gained administrative assistance from the members of the staff in the STEPI who were chosen for this task on the basis of an open competition.

On the other hand, in the UK exercise, both internal manpower from the Assessment and Policy Unit (APU) and the Finance Group (FG), and external manpower from academia, business, related government organisations and, where appropriate, from overseas were involved in the overall process of the evaluation. In particular, several members of the staff of government organisations which have an interest in the research areas of the BBSRC, including the DEFRA, the Food Standard Agency and the Scottish Environment and Rural Affairs Department (SEERAD), played an active role in the expert panels and visiting groups.

Another distinction that can be mentioned in relation to the evaluation unit in the UK system is that the evaluation unit, the APU, carried out two missions during the evaluation process : as the actual evaluator of some elements of the QS and as the secretariat in the 2001 IAE. Meanwhile, in the Korean system, the EMT just organised the evaluation work and contracted out the role of the secretariat to the evaluation committee to an outside institution, the STEPI.

#### ***6.4.5.3. Evaluation Performance***

A comparison of the three evaluation components comprising the evaluation performance, of research institutes in the UK and Korea is summarised in <Table 6-18>.

It seems that there are some differences in the evaluator and the evaluation method, the two components of evaluation performance. Firstly, there is a distinction concerning the composition of the evaluators. As in other evaluation cases, the UK exercise was conducted by mixture of both internal and external experts, whereas the Korean exercise was carried out only by external evaluators. One particular feature concerning the evaluators in the Korean exercise is that they already had relatively high expertise in both R&D matters and internal managerial matters of individual institutes. This was because most of them came from the Planning and Evaluation Committee of the KOCL, a standing

committee entirely made up of external specialists, whose major missions are to coordinate the research functions among member institutes and to provide the Councils with various advisory services.

**<Table 6-18> Comparison of Evaluation Performance for Research Institutes in the UK and Korea**

UK Case	Components	Korean Case
<ul style="list-style-type: none"> <li>- internal evaluators (some QC, CF) <ul style="list-style-type: none"> <li>• staff of APU and FG</li> </ul> </li> <li>- external evaluators (QC, KT, VG) <ul style="list-style-type: none"> <li>• expert panels, visiting groups</li> </ul> </li> </ul>	<b>Evaluator</b>	<ul style="list-style-type: none"> <li>- evaluation committee <ul style="list-style-type: none"> <li>• 12 members of planning and evaluation committee in KOCI</li> <li>• 5 outside specialists</li> </ul> </li> </ul>
<ul style="list-style-type: none"> <li>- qualitative methods <ul style="list-style-type: none"> <li>• review of performance report</li> <li>• site visit (3 days at each one)</li> </ul> </li> <li>- quantitative methods <ul style="list-style-type: none"> <li>• bibliometrics, OPIs</li> </ul> </li> </ul>	<b>Evaluation Method</b>	<ul style="list-style-type: none"> <li>- review of performance report</li> <li>- site visit &amp; interview (2 and a half hours at each one)</li> <li>- questionnaire survey</li> <li><i>* grading system adopted</i></li> </ul>
<ul style="list-style-type: none"> <li>- qualitative criteria <ul style="list-style-type: none"> <li>• prospective KT, VG</li> </ul> </li> <li>- quantitative indicators <ul style="list-style-type: none"> <li>• QS, CF, retrospective KT</li> </ul> </li> </ul>	<b>Evaluation Criteria &amp; Indicators</b>	<ul style="list-style-type: none"> <li>- 22 evaluation criteria</li> <li>- 45 evaluation indicators</li> <li><i>* qualitative one mainly adopted</i></li> </ul>

The evaluation method in the UK exercise had certain noteworthy features. In this exercise, much more quantitative evaluation methods, including bibliometrics and output and performance indicators (OPIs), were utilised. In particular, a bibliometric analysis was broadly adopted in order to measure the scientific and technical excellence of published research articles. In addition, as a part of their site visits, the visiting groups spent three days at each institute in order to examine their overall performance. On the other hand, only two and a half hours was spent on the visit to each institute and interviews in the Korean exercise.

Meanwhile, there is a slight similarity between the two exercise in relation to the evaluation criteria and indicators, in that, in both, the qualitative and quantitative ones were adopted, even though they varied depending on the attributes of the sub evaluation objects.

#### 6.4.5.4. Evaluation Utilisation

A comparison of the two evaluation components comprising the evaluation utilisation of research institutes in the UK and Korea is summarised in <Table 6-19>.

**<Table 6-19> Comparison of Evaluation Utilisation for Research Institutes in the UK and Korea**

UK Case	Components	Korean Case
<ul style="list-style-type: none"> <li>- not available to the public</li> <li>- only a summary posted on website</li> </ul>	<b>Evaluation Report</b>	<ul style="list-style-type: none"> <li>- published</li> <li>- available on the website</li> </ul>
<ul style="list-style-type: none"> <li>- instrumental utilisation               <ul style="list-style-type: none"> <li>• allocation of CSG for 2002-06</li> <li>• adjustment of responsive funds</li> <li>• inputs to the development of management system</li> </ul> </li> <li>- conceptual utilisation               <ul style="list-style-type: none"> <li>• reflection in 10-year scientific road map &amp; 5-year strategic plan of the BBSRC</li> </ul> </li> </ul>	<b>Utilisation Type</b>	<ul style="list-style-type: none"> <li>- instrumental utilisation               <ul style="list-style-type: none"> <li>• submitted to PM and MPB</li> <li>• suggesting correction works</li> <li>• reflection in 2002 strategies and plans in each institute</li> </ul> </li> <li>- conceptual utilisation               <ul style="list-style-type: none"> <li>• not considered</li> </ul> </li> </ul>

First of all, one distinctive difference can be identified in relation to the publication of the final evaluation report. The Korean exercise published this report and also posted it on the website of the KOEI, whereas the UK exercise merely posted a few pages of the summary of the evaluation results to the website.

As in other evaluation cases, a distinction can be made in the utilisation type of the evaluation findings. In the UK system, both instrumental and conceptual utilisation was adopted, even though more emphasis was placed on the former. On the other hand, in the Korean system, the principal type of evaluation utilisation was instrumental, with conceptual utilisation not even being considered.

#### 6.4.6. Comparison of the Evaluation System for R&D Projects

“The review of the Gastrointestinal Response to Food (GRF) programme by the BBSRC in 2000” was selected as the case study of the evaluation exercise of research project in the UK, whereas “the evaluation of the Creative Research Initiatives (CRI) by the MOST

in 2000" was adopted in Korea. In what follows, a similar discussion structure to that for other evaluation exercises will be used in order to compare the evaluation exercises of the research institutes in the UK and Korea.

#### **6.4.6.1. Evaluation Paradigm**

A comparison of the three evaluation components comprising the evaluation paradigm of research projects in the UK and Korea is summarised in <Table 6-20>.

**<Table 6-20> Comparison of Evaluation Paradigm for Research Projects in the UK and Korea**

<b>UK Case</b>	<b>Components</b>	<b>Korean Case</b>
- accountability - programme improvement	<b>Evaluation Purpose</b>	- accountability - programme improvement
- research projects (31) • completed or running	<b>Evaluation Object</b>	- research projects (27) • completed at first stage
- output evaluation - impact evaluation (where appropriate)	<b>Evaluation Type</b>	- output evaluation - assessment of future strategies

From this, it is apparent that there is much similarity in the overall evaluation paradigm of the two evaluation exercises. Since both the UK and the Korean exercises dealt with specific R&D projects included in research programmes, which had already been completed or were still running, the basic principles of R&D evaluation activities were generally applied.

#### **6.4.6.2. Evaluation Resources**

A comparison of the three evaluation components making up the evaluation resources of research institutes in the UK and Korea is summarised in <Table 6-21>. This comparison shows that there are some differences in relation to the three components of evaluation resources.

Firstly, the Korean exercise has one distinctive feature in relation to manpower, in particular, the external manpower. Unlike in other Korean exercises, many prominent

international scientists from 12 countries were actively involved in the evaluation of the CRI programme.

**<Table 6-21> Comparison of Evaluation Resources for Research Projects in the UK and Korea**

UK Case	Components	Korean Case
<ul style="list-style-type: none"> <li>- internal manpower <ul style="list-style-type: none"> <li>• staff of APU (2)</li> </ul> </li> <li>- external manpower <ul style="list-style-type: none"> <li>• review panel (12) <ul style="list-style-type: none"> <li>→ members of agri-food research committee</li> <li>→ independent members</li> </ul> </li> </ul> </li> </ul>	<b>Evaluation Manpower</b>	<ul style="list-style-type: none"> <li>- internal manpower <ul style="list-style-type: none"> <li>• staff of Division of NRDPM in KISTEP (6)</li> </ul> </li> <li>- external manpower <ul style="list-style-type: none"> <li>• evaluation panel (50)</li> <li>• international evaluator (77)</li> <li>• volunteer reviewer (many)</li> <li>• member of PMC (18)</li> </ul> </li> </ul>
<ul style="list-style-type: none"> <li>- APU in BI &amp; I Group</li> <li>- Agri-Food Branch in ST group</li> </ul>	<b>Evaluation Unit</b>	<ul style="list-style-type: none"> <li>- PC for Space Science and Fundamental Research</li> </ul>
<ul style="list-style-type: none"> <li>- questionnaires from grantholders <ul style="list-style-type: none"> <li>• general, meetings, networking/linkage, funding, outputs, other</li> </ul> </li> <li>- bibliometric analysis</li> <li>- final report form</li> </ul>	<b>Evaluation Information</b>	<ul style="list-style-type: none"> <li>- performance report <ul style="list-style-type: none"> <li>• included articles, international conference etc.</li> <li>• research plan for second stage</li> </ul> </li> </ul> <p><i>* very quantitative information</i></p>

Secondly, there is a similarity concerning the evaluation unit in each exercise. In the Korean exercise, as in the UK exercise, the evaluation unit, namely the Programme Coordinator (PC) for Space Science and Fundamental Research of the KISTEP, participated in part of the actual evaluation process, by joining the members of the evaluation panel as well as the Planning and Management Committee (PMC) for the CRI programme.

Finally, another discrepancy can be identified in relation to evaluation information. In the UK exercise, various information sources were considered. In particular, the information from the questionnaire returned by the principal investigators was fully utilised to gain the participants' view on a range of issues, including the achievement of programme objectives, the BBSRC's management system for the programme, new contracts and collaborations, and further funding applied for and received. It is assumed that these kinds of information should act as useful reference for the development of the following stage of the programme itself. Meanwhile, it is apparent that the Korean exercise required information mainly connected to the output of individual research projects, not to

developments at programme level.

#### 6.4.6.3. Evaluation Performance

A comparison of the three evaluation components comprising the evaluation performance of research projects in the UK and Korea is summarised in <Table 6-22>.

**<Table 6-22> Comparison of Evaluation Performance for Research Projects in the UK and Korea**

UK Case	Components	Korean Case
<ul style="list-style-type: none"> <li>- internal evaluator : staff of APU <ul style="list-style-type: none"> <li>• analysis of questionnaire &amp; bibliometrics</li> </ul> </li> <li>- external evaluator : review panel <ul style="list-style-type: none"> <li>• review of APU's work</li> </ul> </li> </ul>	<b>Evaluator</b>	<ul style="list-style-type: none"> <li>- evaluation panel <ul style="list-style-type: none"> <li>• PMC members, related PCs, private sector experts</li> </ul> </li> <li>- international evaluator <ul style="list-style-type: none"> <li>• 3 scientists for each project</li> </ul> </li> </ul>
<ul style="list-style-type: none"> <li>- analysis of grantholder' view from questionnaire</li> <li>- bibliometric analysis of research outputs</li> </ul>	<b>Evaluation Method</b>	<ul style="list-style-type: none"> <li>- analysis of performance report</li> <li>- bibliometric analysis of research outputs</li> </ul>
<ul style="list-style-type: none"> <li>- main evaluation criteria <ul style="list-style-type: none"> <li>• achievement of objectives</li> <li>• scientific excellence</li> <li>• practical significance</li> <li>• improvement of life quality</li> <li>• contribution to UK industrial competitiveness</li> </ul> </li> <li>- various type of evaluation indicators</li> </ul>	<b>Evaluation Criteria &amp; Indicators</b>	<ul style="list-style-type: none"> <li>- 4 evaluation criteria <ul style="list-style-type: none"> <li>• achievement of objectives (20%)</li> <li>• originality of outputs (40%)</li> <li>• potential as world-class researcher (20%)</li> <li>• appropriateness of future plan (20%)</li> </ul> </li> <li>- 5 evaluation indicator</li> </ul>

Regarding the evaluation performance, first of all, there is one striking difference in relation to the role of the evaluator in the two exercises. In the UK system, even though both internal and external evaluators were involved in the actual evaluation process, the internal evaluators, namely the staff of the APU, played a more important role. This is because these staff directly conducted a questionnaire analysis and a bibliometric analysis that were adopted as the main evaluation methods in this exercise, whereas the review panel considered and reviewed the results of the APU's work. As a result, it is clear that the internal evaluators produced more essential information for this evaluation. On the other hand, in the Korean system, the external evaluators, namely international evaluators and the evaluation panels that were entirely made up of private sector professionals, had

the principal responsibility for the actual evaluation.

Even though the two evaluation exercises adopted mainly qualitative criteria, they appealed a different system of weighing to those criteria. While the UK system in general applied the same weight to each evaluation criterion, the Korean exercise applied a different weight to individual evaluation criteria.

#### 6.4.6.4. Evaluation Utilisation

The comparison of the two evaluation components comprising the evaluation utilisation of research project in the UK and Korea is summarised in <Table 6-23>.

**<Table 6-23> Comparison of Evaluation Utilisation for Research Projects in the UK and Korea**

UK Case	Components	Korean Case
<ul style="list-style-type: none"> <li>- not published</li> <li>- circulated to BBSRC office &amp; related staff</li> </ul>	<b>Evaluation Report</b>	<ul style="list-style-type: none"> <li>- not published</li> <li>- circulated to related staff of MOST and KISTEP</li> </ul>
<ul style="list-style-type: none"> <li>- instrumental utilisation <ul style="list-style-type: none"> <li>• decision to continue, reduce or terminate</li> </ul> </li> <li>- conceptual utilisation <ul style="list-style-type: none"> <li>• identification of priority areas for future funding</li> </ul> </li> </ul>	<b>Utilisation Type</b>	<ul style="list-style-type: none"> <li>- instrumental utilisation <ul style="list-style-type: none"> <li>• decision making for further funding → 5 projects terminated</li> </ul> </li> <li>- conceptual utilisation <ul style="list-style-type: none"> <li>• improvement of the future structure of the programme</li> </ul> </li> </ul>

On the basis of this comparison, it is evident that there is a great of similarity between the two cases in relation to the two components of evaluation utilisation. Firstly, both the UK and the Korean exercises did not publish their final evaluation report and circulated it only to a few members of the staff in the related organisations. Then, the evaluation findings of the two exercises were utilised both instrumentally and conceptually, although much emphasis was placed on instrumental utilisation, including decision making for further research funding of research projects.



## 6.4.7. Synthesis

### 6.4.7.1. Section Summary

A comprehensive comparison of the general structure of the national R&D evaluation system in the UK and Korea is shown in <Table 6-24> and comprehensive comparisons of the five evaluation exercises of the two countries are shown in <Table 6-25> to <Table 6-29>. In the following tables, the extent and the reason for the differences between each component in the two exercises are given briefly, on the basis of the basic theoretical background to both policy evaluation and R&D evaluation.

**<Table 6-24> Comprehensive Comparison of Structure of R&D Evaluation Systems in the UK and Korea**

UK Case	Components	Korean Case
five types	Current Type <i>not different</i>	five types
closer relationship	Relations <i>slightly different</i>	weaker relationship
at longer intervals (over two years)	Timing <i>considerably different</i>	at shorter intervals (every year)

**<Table 6-25> Comprehensive Comparison of Evaluation System for S&T and R&D Policy**

UK Case	Evaluation Components		Korean Case
-	Purpose	<i>not different</i>	-
policies & programmes of many departments	Object	<i>considerably different</i>	sub programmes & projects of one ministry
-	Type	<i>not different</i>	-
government staff involved in external manpower	Manpower	<i>slightly different</i>	no government staff involved in external manpower
-	Unit	<i>not different</i>	-
-	Information	<i>slightly different</i>	self-evaluation report mainly used
mixed internal and external evaluators	Evaluator	<i>slightly different</i>	only external evaluator
-	Method	<i>slightly different</i>	meta-evaluative approach adopted
-	Criteria/Indicators	<i>not different</i>	-
-	Report	<i>not different</i>	-
both instrumental & conceptual utilisation	Type	<i>slightly different</i>	only instrumental utilisation

**<Table 6-26> Comprehensive Comparison of Evaluation System  
for S&T or R&D Programme**

UK Case	Evaluation Components		Korean Case
-	Purpose	<i>not different</i>	-
programmes & related policy	Object	<i>considerably different</i>	sub programmes & sub projects
output & impact evaluation	Type	<i>slightly different</i>	mainly interim evaluation
-	Manpower	<i>not different</i>	-
both evaluator and secretariat	Unit	<i>considerably different</i>	only secretariat
various & objective sources	Information	<i>slightly different</i>	simple & subjective sources
internal evaluator	Evaluator	<i>considerably different</i>	external evaluator
various methods adopted	Method	<i>slightly different</i>	mainly literature review
-	Criteria/Indicators	<i>not different</i>	-
-	Report	<i>not different</i>	-
both instrumental & conceptual utilisation	Type	<i>slightly different</i>	only instrumental utilisation

**<Table 6-27> Comprehensive Comparison of Evaluation System  
for Research Councils**

UK Case	Evaluation Components		Korean Case
-	Purpose	<i>not different</i>	-
both strategic & managerial matters	Object	<i>considerably different</i>	managerial matters
-	Type	<i>not different</i>	-
government staff involved in external manpower	Manpower	<i>slightly different</i>	no government staff involved in external manpower
-	Unit	<i>not different</i>	-
various sources used	Information	<i>slightly different</i>	performance report mainly used
mixed internal and external evaluators	Evaluator	<i>slightly different</i>	external evaluators
various methods adopted	Method	<i>slightly different</i>	Mainly literature review
-	Criteria/Indicators	<i>not different</i>	-
-	Report	<i>not different</i>	-
both instrumental & conceptual utilisation	Type	<i>slightly different</i>	only instrumental utilisation

**<Table 6-28> Comprehensive Comparison of Evaluation System  
for Research Institutes**

UK Case	Evaluation Components		Korean Case
-	Purpose	<i>not different</i>	-
four detailed objects	Object	<i>slightly different</i>	more broadly established
-	Type	<i>not different</i>	-
government staff in external manpower	Manpower	<i>slightly different</i>	no government staff in external manpower
both evaluator & secretariat	Unit	<i>considerably different</i>	only organiser
-	Information	<i>not different</i>	-
mixed internal and external evaluators	Evaluator	<i>slightly different</i>	external evaluators
bibliometrics emphasised	Method	<i>slightly different</i>	mainly literature review adopted
-	Criteria/Indicators	<i>not different</i>	-
not published	Report	<i>considerably different</i>	open to the public
both instrumental & conceptual utilisation	Type	<i>slightly different</i>	only instrumental utilisation

**<Table 6-29> Comprehensive Comparison of Evaluation System for R&D Projects**

UK Case	Evaluation Components		Korean Case
-	Purpose	<i>not different</i>	-
-	Object	<i>not different</i>	-
-	Type	<i>not different</i>	-
-	Manpower	<i>slightly different</i>	many international scientists involved
-	Unit	<i>not different</i>	-
questionnaires from researchers mainly used	Information	<i>slightly different</i>	performance report mainly used
mixed internal and external evaluators	Evaluator	<i>slightly different</i>	external evaluators
-	Method	<i>not different</i>	-
same weight to criteria	Criteria/Indicators	<i>slightly different</i>	different weight to criteria
-	Report	<i>not different</i>	-
-	Type	<i>not different</i>	-

It is thus evident that there is no difference in the current types of sub systems of the UK and the Korean national R&D evaluation system. However, there is a considerable difference between these systems in the evaluation timing and a slight difference between

them in the relationship among the individual sub evaluation systems. It is assumed that, in principle, these differences between the two systems have resulted from the respective political and administrative structures as a whole, as well as from the general S&T policies. It is believed that the differences in the relations among the sub R&D evaluation systems have resulted mainly from the different administrative structures of the two countries, whereas the differences in evaluation timing have resulted from their different overall S&T policies.

In the UK, in principle, HM Treasury has the general task of monitoring the effectiveness of all government policies and programmes. As a part of this mission, it conducts a cross-cutting review of science and research every two years. This consists of a spending review to thoroughly examine individual departmental aims and objectives (Jones et al., 2001, HM Treasury, 2000b). According to this basic policy, the evaluation findings of other lower level exercises are expected to be directly or indirectly utilised as the essential inputs to the DTI's spending review and the cross-cutting review of science and research. As a result, the UK R&D evaluation system is categorised as a centralised system (Georghiou, 1998 : 47).

Unlike the UK one, the current Korean R&D evaluation system can be classified as a dispersed system, even though, in general, most S&T related policies are centrally administered by government ministries, such as the NSTC and the MOST. Thus, in Korea, the evaluation systems of major policy issues and of national R&D programmes, which are regarded as the highest level of the national R&D evaluation system, are implemented separately by the OPM/the OGPC and the NSTC, without any communications between them. Therefore, it is apparent that as well as a lack of collaboration during the evaluation process, there is no close relationship in utilising the evaluation findings.

Regarding the differences of evaluation timing, the UK government has a long history of providing the related SET communities with as much autonomy of management and predictability in their research activities as possible. So, it is believed that this policy has led to a longer interval between evaluations, in particular, in the evaluation system for research councils and research institutes. On the other hand, it seems that the Korean S&T policies have placed much more emphasis on the effectiveness of the R&D management system in the public sector than on quality of the research activities themselves. This

policy has resulted in relatively short intervals between evaluations, in particular, in the evaluation system for research councils and research institutes.

Now, the comprehensive comparison of the five detailed evaluation exercises between the UK and Korea will be discussed. Regarding the five sub evaluation exercises, there are many differences in the evaluation systems of the S&T or R&D programmes of the two countries, whereas their evaluation systems for R&D projects have many similarities. On the other hand, in the case of the other three evaluation systems, namely the evaluation systems for S&T and R&D policy, for research councils and for research institutes, there are relatively few differences between the two countries.

Regarding the components of each evaluation system, firstly, many differences between the UK and the Korean exercise can be identified in relation to the following four sub components : the evaluation object, the evaluation manpower, the evaluators and the evaluation unit. On the other hand, there are no differences in relation to another four sub components : the evaluation purpose, the evaluation type, the evaluation criteria & indicators and the evaluation report. The rest of the sub evaluation components, namely, the evaluation information, evaluation method and utilisation type, show a few differences between the two cases.

As with the differences between the UK and Korea in the general structure of their national R&D evaluation systems, it is clear that these differences in the sub components have resulted in part from the different general S&T policies and in part from differences in the development of policy evaluation in general. While the considerable differences in the evaluation object, the evaluation manpower, evaluators and the evaluation unit between the UK and Korea have a close connection with the differences in their general S&T policy, the slight differences in evaluation information, evaluation methods and the utilisation types have to do with the fact that policy evaluation exercises, in particular, the R&D evaluation exercise, are in a developmental stage.

As already discussed, in the UK, as a part of the "customer-contractor principle", the staff of the related government organisations actively participate in the actual evaluation process. However, in Korea, all actual evaluation exercises are conducted by external evaluators from the private sector, even though the internal manpower of the related

government organisations are involved in the evaluation planning stage. This situation has a close relation with the Korean government policy to secure as much fairness and objectiveness in the evaluation exercise as possible. As a result, it is clear that the general S&T policies of the two countries have resulted in considerable differences in some evaluation components, including the evaluation object, evaluation manpower, the evaluators and the evaluation unit.

In addition, it is apparent that the different stage of the development of the evaluation systems and their actual exercise must have affected some of the methodological evaluation components, for example the evaluation information, the evaluation methods and the utilisation type. While the UK government has conducted a policy evaluation and an R&D evaluation since the early 1970s, the Korean government just started to conduct a policy evaluation and an R&D evaluation at national level in the late 1990s. As a result, the UK R&D exercises adopt more varied information sources and more deliberate and complicated evaluation methods, whereas the Korean exercises tend to utilise a single information source and somewhat simplified evaluation methods.

#### **6.4.7.2. Conclusion**

In consequence, it is clear that the differences between the UK and the Korean national R&D evaluation systems are caused in part by difference in administrative structure, in general S&T policies and in the stage of development that policy evaluation has reached in each country. Of these three factors, it is evident that there is a significant difference in the overall political structure of the central government of the two countries in that the UK adopts a parliamentary government system, whereas Korea basically has a presidential system. However, there has been a gradually continuous reduction in the differences between the two countries' administrative structures and government policy in relation to specific issues at managerial level. This is because there has been much policy transfer and mutual lesson drawing among countries throughout the world as a result of various bilateral and multilateral international cooperative activities (Dolowitz et al., 2000).

In particular, the S&T cooperation between the UK and Korea has been fostered by the UK-Korea S&T Cooperation Agreement of 1985 as well as the annual UK-Korea Round

Table Meetings on S&T cooperation, which has been held since 1996 (MOST, 2002i). Through this bilateral cooperation, the Korean government has received useful reference information on S&T governance structure. An example of this is the information it has recently gained on the research council system in public sector R&D management (Yi, 2002 : 9-11). As a result, it is assumed that there are some similarities between the UK and Korea in their S&T related governance structures and their specific policies, even though the political systems upon which their detailed administrative structure is based are completely different.

Meanwhile, it is evident currently that the UK has the most developed national R&D evaluation system of all the European countries, and that this system has had direct or indirect influences on some developing countries, including Korea, as well as on other developed countries (Hills and Dale, 1995 : 35 ; Lee et al., 1994a).

In conclusion, it is believed that there is some possibility of applying elements of the UK R&D evaluation system and its exercise to the Korean case, provided that some minor amendments can be made in the Korean administrative structure and government policies in relation to S&T affairs and that some elements of the methodology it adopts in policy evaluation and R&D evaluation can be improved. Therefore, in accordance with this conclusion, I shall suggest the policy alternatives for the applicability of elements of the UK system to the Korean case in the next chapter.

# **CHAPTER 7. SUGGESTIONS FOR APPLICABILITY OF THE UK CASE TO THE KOREAN CASE**

## **7. 1. Introduction**

From the meta-evaluations and the comparisons between the UK and the Korean national R&D evaluation systems in the previous chapter, it is evident that there are many differences in the fundamental and directive components, including the relationship among the sub evaluation systems, the evaluation timing and the evaluation paradigm. On the other hand, there are relatively few differences in the methodological and instrumental components, for example, evaluation resources, the evaluation performance and the evaluation utilisation. It has been also revealed that the differences between them have resulted in part from the difference in administrative structure and general S&T policies between the two countries and in part from the stage of development that the policy evaluation system, the including R&D evaluation system, has reached in each country.

Despite these differences between the two countries' evaluation systems, I have confirmed that there are possibilities of applying some elements of the UK R&D evaluation system and its exercise to the Korean case. As a precondition of this applying, I also suggested that the Korean governance structures and government policies in relation to S&T affairs need to be slightly adjusted and that some practical elements of both policy evaluation and R&D evaluation also need to be appropriately improved. In what follows, on the basis of the results of comparison of the two national R&D evaluation systems, I shall discuss the policy alternatives for the applicability of elements of the UK system to the Korean case.

## **7. 2. Suggestions for Applicability of the UK System to the Korean System**

In this section, the emphasis of the discussion will be on the evaluation components in



relation to which there are considerable or, at least slight, differences between the UK and the Korean exercises. As discussed in the previous chapter, the meta-evaluation of and the comparison of the components that show such differences indicate that the UK exercise is better than the Korean case in relation to some components, whereas the reverse is the case in relation to others. Meanwhile, there are some components whose strengths and limitations are difficult to discuss in that they normally reflect the differentiated situations or the traditional contexts of individual countries. Thus, these components can be categorised as 'neutral', rather than 'better' or 'worse' in either case. Therefore, some components believed to be neutral or worse in the UK exercise than in the Korean one are excluded from this discussion, even though there are considerable differences in relation to them between the two countries' exercises. This is because my thesis aims not to simply compare the two countries' systems but to draw on useful references for the development of the Korean R&D evaluation systems through the comparative work.

In what follows, I shall discuss the applicability of the UK evaluation exercises to the Korean cases, on the basis of whether particular differences can be categorised as neutral or in terms of one case being better or worse than the other one.

### ***7.2.1. Discussion of the Structure of National R&D Evaluation System***

As already indicated, there is a difference between the UK and Korea in two out of the three sub components of the structure of their national R&D evaluation systems, namely, in the relations among the sub evaluation exercises and in the evaluation timing, in relation to both of which the UK system is superior to the Korean one. Therefore, further discussion will be confined to these two evaluation components.

Unlike in the UK evaluation exercises in which there are relatively close relationships among the sub evaluation exercises in the national R&D evaluation system, there has been a criticism of the lack of close linkages among the sub evaluation exercises making up the Korean national R&D evaluation system. Therefore, it appears that the Korean government can obtain useful information from the UK case to strengthen the mutual relationships among the sub evaluation exercises of its national R&D evaluation system.

First of all, the relations between the evaluation of major policy issues and other

evaluation exercises needs to be strengthened both at the direct and indirect levels, given that the evaluation system of the major policy issues, in particular of S&T related policy issues, can be improved to deal with the cross-departmental policies and strategies of the central government ministries interested in S&T and R&D matters. In particular, when the above improvements have been made, the findings of the evaluation of both the national R&D programmes by the NSTC and of the research councils by the OGPC should be key inputs to the process of the evaluation of the major policy issues in S&T areas. This is because these two evaluation exercises examine most R&D programmes sponsored by the government ministries as well as by the research councils that have the principal responsibility for the management of the GRIs who are main public sector R&D performers in Korea.

In addition, the mutual relationship between the evaluation system for national R&D programmes and that for the research institutes under the three S&T related research councils needs to be strengthened. Even though these two evaluation systems deal mainly with the R&D activities implemented by many GRIs, they are separately conducted from the perspective of the needs of two individual central government organisations, namely the NSTC and the OGPC. The fact that the evaluation exercise is duplicated by the two systems is considered as a cause of administrative burden on both the research institutes and researchers. Therefore, it is necessary that the findings of the two evaluation exercises, in particular about R&D matters, should be shared.

The adjustment of the evaluation timing, namely the intervals between regular evaluation exercises, will be now discussed. Relatively longer intervals at which the UK exercises are conducted can be an excellent reference for the Korean cases. As already noted, four out of the five Korean R&D evaluation exercises, i.e. excluding the evaluation system for the CRI programme, are conducted on an annual or a six monthly basis. I have also pointed out that this shorter interval between each evaluation exercise has resulted mainly in the establishment of a shorter evaluation paradigm, including, the evaluation purposes and the evaluation object, as well as in the instrumental utilisation of the evaluation findings. Therefore, the evaluation timing of the four evaluation exercises should be extended in accordance with the attributes of the individual evaluation systems, which are expected to function properly in the overall national R&D evaluation system.

Firstly, the intervals between the evaluation exercises in relation to major policy issues need to be extended from the current six months to two or three years. This is because the effects and influences of a certain policy tend to be realised some years after its implementation or completion, whether they are short-term and direct or long-term and indirect. For this reason, there is a strong need for the evaluation of the S&T related policy issues to be conducted as an output evaluation and an impact evaluation from a relatively long term perspective, rather than as an interim evaluation of policy implementation from a short term perspective.

Secondly, it is necessary that the interval between evaluations of national R&D programmes should be extended from a year to two or three years. In general, there has been a tendency for their whole effect or influence of large-scale R&D programmes to be produced in some period after their completion.

Thirdly, there is a strong need to extend the intervals between the evaluations of both the research councils and the research institutes. In both cases, the current evaluation system is conducted every year so that it places more emphasis on the effectiveness of the management at individual institutions' operating levels than on the rationality and the appropriateness of the developmental and strategic matters of S&T governance in the public sector. Therefore, the intervals between the evaluation exercises of both systems should be extended from a year to three or four years. In this case, extended intervals between evaluation exercises are expected to be compatible with, for example, the period of the service of the chairpersons of the research councils and the presidents of the research institutes. In particular, since the Korean GRIs will continue to play a principal role in public sector R&D activities, there is a strong requirement for those R&D activities to be allocated to them on a more stable and predictable basis by rationally extending intervals between evaluation exercises.

### ***7.2.2. Discussion of the Evaluation System for S&T and R&D Policy***

As already discussed, there are 11 components in both the UK and the Korean systems of evaluation of S&T and R&D policy. There is a slight difference between the two systems in five of these components and a considerable difference in relation to one of the component. Of these six components, two are believed to be better in the UK exercise,

whereas four are considered as neutral between the two exercises or better in the Korean one. A brief explanation of these components that are neutral or better in the Korean exercise will be given and then the discussion will continue in relation to how aspects of the components that are believed to be better in the UK exercise can be applied to the Korean case.

Both the evaluation manpower and the evaluators can be classified as components that are neutral between the two systems. In fact, they usually result from the differentiated S&T policies whose strengths and weaknesses are difficult to judge at the moment. It is assumed that the UK government places more emphasis on the relevance and the acceptance of the evaluation findings than on the fairness of the evaluation process by making the related staff of the government departments actively participate in the actual evaluation exercise, whereas the Korean government aims to secure the objectiveness and fairness of the evaluation process by having it conducted only by external evaluation manpower entirely from the private sector, including the universities, industry, the research institutes and the media. In particular, this principle of the Korean government in relation to evaluation manpower and the evaluators applies equally to other four exercises making up the national R&D evaluation system.

Meanwhile, it is judged that the Korean exercise is slightly better than the UK case in regard to the evaluation information and the evaluation method. It adopts the meta-evaluative approach as a main evaluation method so that it aims to enhance the capability of the evaluation activities of the central government ministries.

As a result, there are some possibilities of applying elements of the UK exercise relating to both the evaluation object and the utilisation type of evaluation findings to the Korean case.

Firstly, it is clear that the establishment of the evaluation objects in the UK exercise can be an excellent reference for the Korean case. This is because, through the meta-evaluation, it has been confirmed that there is a strong need to adjust the evaluation objects of the major policy issues, in particular, the S&T related policy issues, in the Korean system. Unlike in the UK exercise, where both the SET policies and their main programmes implemented by the central government departments as well as by their

main NDPBs are involved, the Korean case includes only sub programmes and main projects comprising the pertinent policy conducted by only one central ministry, for example, the MOST. So, there are some limitations in that the Korean case cannot evaluate the pertinent policy itself as a whole and also cannot manage to handle the cross-departmental S&T and R&D policies at national level. Therefore, the objects in the evaluation of the S&T and R&D policies should be adjusted in order to deal with as many related policies in the central administrative offices as possible, including the MOE, the MOCIE, the MIC and the MOHW.

Secondly, it appears that the UK exercise can be good example for the Korean case in relation to the improvement of utilisation types of the evaluation findings. Currently, while both instrumental and conceptual utilisations are equally adopted in the UK system, greater emphasis is placed on instrumental utilisation in the Korean system. Therefore, it is necessary that the utilisation type of the evaluation findings in the Korean exercise should be extended in accordance with the improved evaluation paradigm, in particular, the amendment of the evaluation objects of the S&T related policy issues.

### ***7.2.3. Discussion of the Evaluation System for S&T or R&D Programme***

There are 11 components in both the UK and the Korean systems of the evaluation of the S&T or R&D programmes. There is a slight difference between the two national evaluation systems in three of the components and a considerable difference between them in relation to four components. Of these seven components, five are believed to be better in the UK exercise, whereas two are regarded as neutral between the two countries. A brief explanation of the neutral components will now be given and then a discussion will continue in relation to how aspects of the components that are believed to be superior in the UK exercise can be applied to the Korean case.

Both the evaluation unit and the evaluators can be classified as neutral components, even though there are considerable differences between the UK and the Korean exercises. As I have already explained why the evaluators can be classified as components that are neutral between the two countries' cases, further discussion will be related to the evaluation unit. It is evident that the difference in role of the evaluation unit has arisen due to the general policy in relation to the actual evaluation exercise. While the UK

government, in particular the DTI, has a policy to secure as much relevance and professionalism in the evaluation exercise as possible by ensuring that a high expertise evaluation unit participates in the actual evaluation process, the Korean government, in the form of the NSTC and the MOST, places an emphasis on the objectiveness of the evaluation exercise by excluding the evaluation unit from the actual evaluation exercise. Hence, in most Korean evaluation exercises, the role of the evaluation unit is confined to acting as the secretariat to the evaluation committee or the organiser of the evaluation work.

Consequently, there are some possibilities of applying aspects of the following four components of the UK exercise to the Korean case : evaluation objects, evaluation type, evaluation methods and utilisation type.

First, it is evident that the establishment of the evaluation objects in the UK exercise can act as very useful reference to the Korean case. The evaluation objects in the UK exercise were set up at the level of programmes and, where appropriate, a certain policy, whereas in the Korean exercise, the evaluation objects were determined at the level of sub R&D programmes and projects. For this reason, the Korean system has a deficiency in not examining the overall performance of the pertinent R&D programme, but rather investigating the partial and short-term outputs of the sub programmes and projects. Therefore, there is a strong need for the level of the evaluation object to be upgraded to that of the programme, provided that the intervals between evaluation exercises are extended from the current one year to two or three years.

Second, it is apparent that the UK exercise can act as good reference for the Korean case in relation to the evaluation type. The UK exercise was conducted as an output evaluation, impact evaluation and assessment of future management, whereas the Korean exercise focused on an interim evaluation and output evaluation. Therefore, there is a strong need for the evaluation type of the Korean exercise to be changed from being an interim evaluation and output evaluation to being an output and an impact evaluation. As already discussed, the limitation of the evaluation types in the Korean system resulted in part from the shorter interval between evaluation exercises and in part from the establishment of the evaluation objects at a lower level. Therefore, it is inevitable that the evaluation types will be diversified in parallel with the extension of the evaluation timing as well as

the improvement of the evaluation objects.

Third, it is expected that the Korean government can utilise the UK exercise as a source of useful information regarding evaluation methods. In order to evaluate the past performance and anticipated future impact of the S&T programme, the UK government adopted diverse evaluation methods, including a field survey, a literature review, an in-depth interview and comparative studies, whereas the Korean exercise relied mainly on a literature review in the form of the review of the performance report submitted by the individual ministries to be evaluated. Therefore, once the above improvements in evaluation objects and evaluation types have been made, the Korean exercise needs to adopt more diverse evaluation methods, for example, field surveys and comparative studies with similar cases in other countries.

Fourth, it is apparent that the UK exercise can provide the Korean case with a good example in term of the utilisation types of evaluation findings. The evaluation findings were utilised from both the instrumental and conceptual perspective in the UK system, whereas greater emphasis was placed on instrumental utilisation in the Korean system. Therefore, it is necessary for the utilisation type of the evaluation findings in the Korean exercise to be extended in accordance with the amendment of the evaluation objects and the diversification of the evaluation types.

#### ***7.2.4. Discussion of the Evaluation System for the Research Councils***

There are 11 components in the UK and the Korean systems of the evaluation of the research councils. There is a slight or considerable difference between the two national evaluation systems in relation to six of these components. Of these six components, four are believed to be better in the UK exercise, whereas two are regarded as neutral between the two countries' exercises. A brief explanation of the neutral components will now be given, followed by a discussion of how the aspects of components which are better in the UK exercise can be transferred to the Korean case.

As in other evaluation exercises, there are slightly differences between the UK and the Korean system in both evaluation manpower and the evaluators. Despite the differences, I already pointed out that these components could be classified as neutral due mainly to the

differentiated S&T policies of the two countries. So, no further explanation will be given regarding these components.

In consequence, there are some possibilities of applying aspects of the three following evaluation components of the UK exercise to the Korean case : evaluation objects, evaluation methods and the utilisation type of evaluation findings.

First, there is a strong need for the Korean exercise to adjust its evaluation objects by taking the UK exercise as a benchmarking model. There is a considerable difference between the UK and the Korean system in evaluation objects due mainly to the evaluation timing. The UK exercise established its evaluation objects in order to deal with fundamental and strategic matters, including the mission, structure and governance of the research council system as well as operating and managerial matters such as the internal management system of the individual research councils. On the other hand, the Korean exercise confined the evaluation objects to managerial matters at operating level. Therefore, in the Korean evaluation system, when the above changes regarding the evaluation timing have been made, the evaluation object should be extended to include strategic and long-term matters seen from the viewpoint of public S&T governance at the national level in addition to current short-term and managerial matters.

Second, it is evident that the UK exercise can provide the Korean case with useful references in relation to the evaluation method. Even though there is a slight difference between the two countries, the UK exercise has a distinctive characteristic in that public consultation has to be carried out in each review stage according to the Cabinet Office's *Guidance on Agency and NDPB Quinquennial Review*. In the UK system, the extensive conduct of the public consultation has been considered as one of the essential evaluation methods to be able to reflect diverse opinions of the key stakeholders, including government, universities, PSREs and business, into the evaluation process. Therefore, it is necessary that the Korean government consider public consultation as an evaluation method, when the above changes regarding the evaluation timing and the evaluation objects have been made. The reason is that the improved evaluation system of the research councils accompanying the extension of the evaluation timing and the amendment of the evaluation objects will have a great influence on the related science and research communities, including the universities, many GRIs and industry.



Third, it is believed that the Korean exercise can gain useful information from the UK case regarding the utilisation types of the evaluation findings. Unlike in the UK exercise where both instrumental and conceptual utilisations were adopted, the evaluation findings were mainly utilised from the instrumental perspective in the Korean system. However, once both the evaluation timing and the evaluation objects have been improved, there is a strong need for utilisation types of the evaluation findings to be extended to conceptual utilisation to handle the long-term and strategic affairs of the S&T governance structure in the public sector, including the review of the organisational status of the research council, relations between the research councils and the key stakeholders, and their internal management systems.

#### ***7.2.5. Discussion of the Evaluation System for the Research Institutes***

There are 11 components in both the UK and the Korean systems of the evaluation of the research institutes under the responsibility of the research councils. There is a considerable difference between the two national evaluation systems in two of these components and a slight difference between them in five of these components. The UK exercise is believed to be superior in three of these seven components, whereas four are regarded as neutral between the two exercises or better in the Korean exercise. A brief explanation of the neutral components will now be given, followed by a discussion of how the aspects of these components that are better in the UK exercise can be transferred to the Korean case.

As in other evaluation exercises, there are slightly differences between the UK and the Korean system in both evaluation manpower and the evaluators. In addition, there is a considerable difference in the evaluation unit in that the evaluation unit acted as both internal evaluators and secretariat to the expert panel in the UK exercise, whereas it acted as only organiser of the evaluation work in the Korean exercise. Despite the differences of the three components, further discussion will not be given. This is because the differences of these components came mainly from the differentiated S&T policies of the two countries.

Meanwhile, it is thought that the Korean exercise is superior to the UK case in its dissemination of the evaluation report. The UK exercise did not publish the final

evaluation report and merely posted the summary of the result on the website, even though it is expected to function as an essential reference for the related research institutes as well as for the research communities concerned.

As a result, there is possibility of applying the aspects of the following components of the UK exercise to the Korean case : evaluation objects, evaluation methods and the utilisation type.

First, it is desirable that the Korean exercise should adjust the evaluation objects by adopting the UK case as a useful reference. It is believed that the UK system is better than the Korean one in the detailed contents and attributes of the evaluation objects, even though the two countries proposed research matters and internal managerial matters as the main evaluation objects. While the UK exercise placed greater emphasis on core research outputs and R&D activities than on the R&D management system, the Korean exercise focused more on R&D managerial matters than the research activities and their quality. Therefore, it is essential that the Korean exercise should re-establish the evaluation object to properly deal with the research activities and their quality as well as the R&D management system.

Second, the UK exercise is expected to provide a good reference for the Korean exercise in relation to the evaluation method. In particular, a bibliometric analysis was broadly adopted in order to measure the scientific and technical excellence of published research articles in the UK system. Thus, in addition to the current evaluation methods, the Korean system needs to make greater use of various evaluation methods to judge, in particular, the quality of the research outputs and to ensure that they reflect the amended evaluation objects.

Third, it is evident that the UK exercise can provide a good reference to the Korean case in relation to the utilisation types of the evaluation findings. The UK exercise utilised the evaluation findings to allocate research funds to the individual research institutes under its responsibility as well as to develop the council's ten-year scientific opportunities road map and the next five-year strategic research plan. On the other hand, the Korean exercise mainly utilised them to adjust the funding allocation and to suggest the correction works. Therefore, in the Korean system, there is a strong need to extend the utilisation types of

the evaluation findings to cover the conceptual perspective at research council level as well as the instrumental perspective at member research institute level, once the evaluation timing have been properly extended and the evaluation objects have been rationally revised.

#### ***7.2.6. Discussion of the Evaluation System for the R&D Projects***

There are slight differences between the UK and Korean evaluation systems of national R&D projects in relation to four of the 11 evaluation components. Of these, three components are regarded as neutral between the two systems, whereas the Korean exercise is superior in relation to one component. Consequently, there is no possibility of applying aspects of the components of the UK exercise to the Korean case. A brief discussion will now be devoted to the components that are neutral between the two countries' systems or better in the Korean case.

As in other evaluation exercises, there are slight differences between the UK and the Korean system in three components : evaluation information, evaluators and evaluation criteria. However, it is evident that these differences between the two countries mainly arose from the differentiated S&T policies and the general evaluation system of each country, rather from the strengths or limitations of one country compared to another. Therefore, it is thought currently that any discussion of the applicability of elements of the UK exercise to the Korean case would have no practical implications.

Meanwhile, it is believed that the Korean exercise is better than the UK case in the composition of its evaluation manpower. Unlike in the UK exercise, many prominent international scientists from 12 countries were actively involved in the Korean exercise. The active involvement of prominent international evaluators is considered one of the practical methods to secure the expertise and high quality of the evaluation findings, bearing in mind that most national R&D projects usually deal with leading edge science and technology.

### 7.3. Chapter Conclusions

As a concluding remark, it is evident that, among the five evaluation exercises in the UK, the evaluation system of the S&T programme can provide many useful references for the future development of the Korean system, because the four evaluation components of the UK exercise, namely evaluation objects, evaluation types, evaluation methods and the utilisation type, can function as benchmarks for the Korean system. In addition, there is possibility of applying aspects of the UK exercise to the Korean evaluation system for both the research councils and the research institutes. In this case, it is clear that three components of the UK evaluation exercises, namely evaluation objects, evaluation methods and the utilisation type, can be adopted as good practices by the Korean government. However, it is believed that there is no possibility of the Korean evaluation system of R&D projects adopting the practices of the UK exercise.

Meanwhile, among individual evaluation components, the evaluation objects of the UK exercises are expected to be the best reference for the Korean cases. Therefore, it is clear that four out of the five Korean evaluation exercises, i.e. excluding the evaluation system of the national R&D project, can re-establish their individual evaluation objects so that they relate to more fundamental and strategic matters by taking the UK exercise as a benchmarking model. In addition, the above four Korean exercises can draw good examples from the UK cases in relation to the utilisation type of the evaluation findings. Regarding evaluation methods, the Korean evaluation systems for the research councils and research institutes may adopt aspects of the UK experience, such as public consultation in the quinquennial review of research councils and the bibliometric analysis in the 2001 IAE.

Finally, as a precondition of the above changes and improvements, I have already suggested that the four Korean R&D evaluation systems, excluding that for the national R&D project, should extend the intervals between evaluation exercises from one year or a six months to longer period, for example, two, three or four years depending on the attributes of the individual R&D evaluation systems. For this purpose, I am convinced that the UK exercises can provide the best references for the Korean exercise.

## CHAPTER 8. CONCLUSIONS

### 8. 1. Synthesis of Research

The major concern of my thesis is to suggest development for the Korean national R&D evaluation system, taking the UK case as a benchmarking model. In order to do this, I firstly conducted the two case studies of the UK and the Korean national R&D evaluation systems each of which covered five evaluation exercises : (1) the evaluation of S&T and R&D policy, (2) the evaluation of S&T or R&D programmes, (3) the evaluation of research councils, (4) the evaluation of research institutes under the research councils and (5) the evaluation of R&D projects. Then, in order to identify the strengths and limitations of the UK and the Korean systems as well as the differences between the two countries in relation to the five evaluation exercises, the meta-evaluations and the comparisons of those meta-evaluations in the two countries were conducted. Lastly, on the basis of the results of the previous work, I discussed the applicability of some elements of the UK system to the Korean case. This discussion focused on some evaluation components which are expected to be better in the UK case than in the Korea one as well as the fact that there are differences between the two countries' systems.

In what follows, I shall first summarise key findings of individual chapters which have dealt directly with the two countries' case studies, their meta-evaluations and comparisons. Then, I shall provide some key suggestions for development of the Korean national R&D evaluation system, which are mainly based on the findings of the meta-evaluations and comparisons.

#### *8.1.1. Summary of the Thesis*

This part will be devoted to summarising contents of the case studies and the meta-evaluations of the UK and the Korean national R&D evaluation systems as well as the comparisons between those meta-evaluations.

#### ***8.1.1.1. Summary of the UK Case Study and Meta-evaluation***

To carry out the case study and the meta-evaluation of the UK national R&D evaluation system, the following five evaluation exercises were selected : (1) the cross-cutting review of science and research in 2000, (2) the quinquennial review of the six grant awarding research councils in 2001, (3) the institute assessment exercise (IAE) by the BBSRC in 2001, (4) the evaluation of funding for UK civil space activities conducted by the DTI in 2001 and (5) the review of the gastrointestinal response to food (GRF) programme conducted by the BBSRC in 2000.

On the basis of the case studies of the five evaluation exercises in the UK S&T and R&D sectors, it seems that, generally in the UK, evaluation exercises in these sectors have similar characteristics, even though they differ slightly in some aspects such as evaluation resources and evaluation performance, according to the evaluation purposes. While external evaluators played major roles in the 2000 cross-cutting review of science and research and the 2001 quinquennial review of the six grant awarding research councils, internal evaluators had more important missions in the evaluation of both the UK civil space activities conducted by the DTI and the GRF programme conducted by the BBSRC, even though they were in collaboration with external evaluators. Also, at the higher level of evaluation exercises such as the 2000 cross-cutting review of science and research and the 2001 quinquennial review of research councils, more qualitative evaluation methods were adopted. However, in the case of evaluation exercises at the project-level such as the GRF programme, the emphasis was placed on quantitative evaluation methods. In particular, it is apparent that, in all the evaluation exercises, the findings of evaluation have been fully reflected in the decision-making processes in relation to future budget allocation and further strategic development of related policy, programmes and institutions.

The principal findings of the meta-evaluation of the five UK evaluation exercises will now be summarised. In principle, it seems that the evaluation resources and the evaluation performance, which are considered as the instrumental and methodological components of the R&D evaluation system, were sufficient and suitable, whereas the evaluation paradigm, which is regarded as a more fundamental and directive component, needs to be improved. Meanwhile, the evaluation utilisation, which is to provide it

feedback process for the next stage of the policy process, was relatively appropriate.

Having outlined the above general findings of the meta-evaluation, the more detailed findings of evaluation of the four individual main evaluation components will now be summarised.

Firstly, in three out of the five evaluation exercises, i.e. excluding both the quinquennial review of research councils and the review of the GRF programme, the evaluation purposes were not rationally established. Therefore, there is a need to adjust them to include knowledge enhancement. In addition, an impact evaluation needs to be added to the three above-mentioned evaluation systems as further evaluation type. While improvements need to be made to the evaluation purposes and types, the definition of the evaluation objects was completely satisfactory. In particular, the evaluation objects of the 2000 cross-cutting review, the 2001 quinquennial review of the six research councils and the evaluation of the UK civil space activities were very properly set up.

Secondly, in general, it is believed that the evaluation resources used were sufficient and suitable, even though a few improvements are still needed in some evaluation exercises. First of all, one feature of the evaluation resources can be extracted from the composition of the evaluation manpower. Of the evaluation manpower, the external manpower comes from academia, industry, the research institutes and government departments concerned. That is, in the UK R&D evaluation exercises, the related government departments normally participated in the overall process of evaluation of their policies and programmes. In particular, in some evaluation exercises, the evaluation unit acted as the secretariat to the evaluation panels as well as being directly involved in the actual evaluation of some sub evaluation objects. This is regarded as one of the distinctive characteristics in the UK R&D evaluation system. Despite the general strengths of the evaluation resources, in the case of the 2000 cross-cutting review, the principal evaluation information came from existing research, which did not include first hand materials but second hand ones. Therefore, much more effort needs to be made in order to get and use evaluation information from a greater variety of sources.

Thirdly, it is thought that the evaluation performance was relatively effectively implemented. In particular, in some evaluation exercises such as the 2001 IAE, the

evaluation of the UK civil space activities and the review of the GRF programme, the internal evaluators played more important roles than the external evaluators did, by preparing the common evaluation protocol for the evaluation of the UK civil space activities and conducting the bibliometric analysis of the 2001 IAE and the evaluation of GRF programme. However, there is also a need for improvement in the evaluation methods, in particular, in the case of the 2000 cross-cutting review. This is because, in this review, a rather simple evaluation method, in the form of a review of the small amount of existing research available, was adopted.

Finally, in principle, it is clear that, in the five evaluation exercises, instrumental and conceptual utilisation of the evaluation findings was excellent. However, there is a need for improvement in relation to the dissemination of the evaluation report of the BBSRC's two exercises in that they were circulated only in the office and were not made available to the public.

Consequently, it is apparent that, among the five UK R&D evaluation systems, both the 2001 quinquennial review of research councils and the evaluation of funding for UK civil space activities are better examples of evaluation exercises in the UK. Meanwhile, the two evaluation exercises conducted by the BBSRC, namely the 2001 IAE and the evaluation of the GRF programme, could be considered as relatively good cases. On the other hand, it seems that there is a need for further improvement in the 2000 cross-cutting review of science and research.

#### ***8.1.1.2. Summary of the Korean Case Study and Meta-evaluation***

For the Korean case study and its meta-evaluation, five evaluation exercises were also selected : (1) the evaluation of major policy issues of the MOST in 2001, (2) the evaluation of national R&D programmes in 2002, (3) the evaluation of the five research councils in 2001, (4) the evaluation of member research institutes conducted by the KOICI in 2002 and (5) the evaluation of the creative research initiatives (CRF) conducted by the MOST in 2000.

From these case studies, it appears that, generally in Korea, evaluation exercises in the S&T and R&D sectors have similar characteristics in most aspects of evaluation



components. Firstly, it is evident that external evaluators had an actual responsibility for the exercise, because all five evaluation exercises were conducted by independent evaluation committees whose members came entirely from industry, academia, the research institutes and the media. Next, budget adjustment or allocation was considered as the most frequent and important use to which the evaluation findings were put. Meanwhile, another distinction of the Korean R&D evaluation system is that all evaluation exercises have been performed every year without consideration of the particular attributes and characteristics of individual evaluation systems. However, in contrast to other Korean cases, the evaluation exercise of R&D projects under the CRI programme represents a recent policy change undertaken by the Korean government to secure the objectiveness and professionalism of the evaluation.

Now, the principal findings of the meta-evaluation of the five Korean evaluation exercises will be summarised. On evaluating the five Korean evaluation exercises, in principle, it is apparent that four of these systems, i.e. excluding the evaluation system for the CRI programme, are at an early stage in their development. Therefore, it seems that there are many improvements that need to be made in the directive and fundamental components of R&D evaluation systems, even though the instrumental and methodological components are relatively sufficient and suitable. In addition, four out of the five exercises, i.e. excluding the CRI case, have a distinctive feature in that they are carried out every year. There is no doubt that such a short interval between evaluation exercises has had a considerable influence on the overall evaluation system and its actual implementation.

Now that the general findings of the meta-evaluations have been outlined, the more detailed findings of the evaluation of the four main evaluation components will be summarised.

Firstly, in four out of the five evaluation exercises, i.e. excluding the evaluation of the CRI programme, the evaluation purposes were established in terms of a relatively narrow perspective. Therefore, there is a strong need that it should be re-established to include a broader and longer term purposes such as the enhancement of basic knowledge. In addition, along with the extension of the evaluation purposes, an impact evaluation needs to be added as further evaluation type to the four evaluation systems other than the

evaluation of the CRI programme.

Meanwhile, there are two essential preconditions that must be met, if these improvements in the evaluation purposes and types are to be realised. First of all, in most Korean evaluation systems, the evaluation object should be re-established not at the short-term and operating level but at the more long-term and strategic level. Next, the related government policies, in particular regarding the research councils and their research institutes, should be transferred from their current emphasis on managerial matters to an emphasis on the systematic and strategic R&D governance structure in the public sector at national level.

Secondly, it seems that the evaluation resources used were sufficient and suitable, bearing in mind that the evaluation paradigm of most Korean evaluation systems was established largely at the operating level rather than at the strategic level. However, it is evident that, in the future, the overall evaluation resources should be changed in accordance with the improved evaluation paradigm discussed above.

Despite the improvements needed for the future, at the current stage, some distinctions can be drawn in relation to the evaluation resources. Firstly, unlike in the UK exercises which some related government officials were involved in, the five Korean evaluation exercises were conducted by external manpower entirely made up of private sector experts from academia, industry, the research institutes and the media. Another distinction can be drawn in relation to the evaluation information, because the evaluators relied mainly on second hand sources such as performance reports submitted by the institutions to be evaluated. There is also one distinction to be made in relation to the evaluation units. In the five Korean evaluation exercises, evaluation units acted only as the secretariat to the evaluation committee or evaluation panel, unlike in the UK evaluation cases, where the members of the evaluation unit were directly involved in the actual evaluation of some sub evaluation objects.

Thirdly, it is thought that the evaluation performance was relatively effectively implemented, because the initial government intentions regarding R&D evaluation exercises were focused on managerial matters in relation to policies, programmes and research institutions at the operating level. In particular, the use of international

evaluators in the evaluation of the CRI programme is considered as a very distinctive feature in the evaluation performance. In addition, the adoption of the meta-evaluative approach in the evaluation of major policy issues is considered as one of the distinctive features of the evaluation methods. However, there is a need for improvement in the evaluation criteria used, in particular, in the evaluation of the member research institutes under the KOEI. This is because too many evaluation criteria and indicators were adopted and this may have placed an unexpected administrative burden on the evaluators as well as on related staff and institutions.

Finally, it is clear that, in the five evaluation exercises, much more emphasis was placed on the instrumental utilisation of the evaluation findings from the short-term and managerial perspectives than on their conceptual utilisation from the long-term and strategic perspectives. Therefore, the evaluation findings of the five R&D evaluation cases are expected to contribute more to the immediate improvement of the related R&D programmes and the management system of the research institutions at ministry level than to the future development of S&T and R&D policies and their governance system at national level.

In consequence, it is apparent that, among the five Korean R&D evaluation systems, the 2000 evaluation of the CRI programme is the best of the examples of evaluation exercises in Korea. Meanwhile, both the evaluation of the national R&D programmes and the evaluation of the member research institutes under the KOEI can be considered as relatively satisfactory cases. On the other hand, it seems that there are a great many improvements required in both the evaluation of major policy issues and the evaluation of the five research councils.

#### ***8.1.1.3. Summary of the Comparison of the UK and the Korean Systems***

In this thesis, the comparison of the UK and the Korean national R&D evaluation systems have been conducted from the following three different, but closely connected viewpoints : (1) a comparison of the general structure of the national R&D evaluation system consisting of sub evaluation systems, their relations and evaluation timing, (2) a comparison of the individual evaluation exercises making up the entire R&D evaluation system and (3) a comparison of individual sub components of each evaluation system.

First of all, a comparison of the general structure of national R&D evaluation systems in the UK and Korea will be briefly given. There is no difference in the current types of sub systems of the UK and the Korean national R&D evaluation system, since the two countries' national R&D evaluation systems comprise the same five sub evaluation systems.

However, there is a considerable difference between these systems in the evaluation timing and a slight difference between them in the relationship among the individual sub evaluation systems. It is assumed that, in principle, these differences between the two systems have resulted from the respective political and administrative structures as a whole, as well as from the general S&T policies. It is believed that the differences in the relations among the sub R&D evaluation systems have resulted mainly from the different administrative structures of the two countries, whereas the differences in evaluation timing have resulted from their different overall S&T policies.

Secondly, the comprehensive comparison of the five detailed evaluation exercises between the UK and Korea will be summarised. Regarding these sub evaluation exercises, there are many differences in the evaluation systems of the S&T or R&D programmes of the two countries, whereas their evaluation systems for R&D projects have many similarities. On the other hand, in the case of the other three evaluation systems, namely the evaluation systems for S&T and R&D policy, for research councils and for research institutes, there are relatively few differences between the two countries.

Finally, the comparison of individual components of each evaluation system reveals many differences between the UK and the Korean exercise in relation to the following four sub components : the evaluation object, the evaluation manpower, the evaluators and the evaluation unit. On the other hand, there are no differences in relation to another four sub components : the evaluation purpose, the evaluation type, the evaluation criteria and indicators, and the evaluation report. The rest of the sub evaluation components, namely, the evaluation information, evaluation method and utilisation type, reveals a few differences between the two cases.

As with the differences between the UK and Korea in the general structure of their national R&D evaluation systems, it is clear that these differences in the sub components

have resulted in part from the different general S&T policies and in part from differences in the development of policy evaluation in general. While the considerable differences between the UK and Korea in evaluation objects, evaluation manpower, the evaluators and evaluation units have a close connection with the differences in their general S&T policy, the slight differences in evaluation information, evaluation methods and the utilisation types have to do with the fact that policy evaluation exercises, in particular, the R&D evaluation exercise, are at a developmental stage.

### ***8.1.2. Key Suggestions for the Applicability of Aspects of the UK Case to the Korean Case***

From the meta-evaluations and comparisons of the UK and the Korean cases, it is evident that there is some possibility of applying elements of the UK R&D evaluation system and its implementation to the Korean case, provided that some minor amendments can be made in the Korean administrative structure and government policies in relation to S&T affairs and that some elements of the methodology it adopts in policy evaluation and R&D evaluation can be improved.

First of all, as a precondition for the changes and improvements discussed below, I suggest that the four Korean R&D evaluation systems, excluding that for R&D projects of the MOST, should extend the intervals between evaluation exercises from one year or six months to a longer period, depending on the attributes of the individual R&D evaluation systems. For this purpose, I am convinced that the UK exercises can provide the best references for the Korean exercise.

Next, it is evident that, among the five evaluation exercises in the UK, the evaluation system of the S&T programme can provide many useful references for the future development of the Korean system, because the four evaluation components of the UK exercise, namely evaluation objects, evaluation types, evaluation methods and the utilisation type, can function as benchmarks for the Korean system. In addition, there is a possibility of applying aspects of the UK exercise to the Korean evaluation system for both the research councils and the research institutes. In this case, it is clear that three components of the UK evaluation exercises, namely the evaluation objects, the evaluation

methods and the utilisation type, can be adopted as good practices by the Korean government. However, it is believed that there is no possibility for the Korean evaluation system of R&D projects to adopt the practices of the UK exercise.

Lastly, among individual evaluation components, the evaluation objects of the UK exercises are expected to be the best reference for the Korean cases. Therefore, it is clear that four out of the five Korean evaluation exercises, i.e. excluding the evaluation system for the national R&D project, can revise their individual evaluation objects so that they relate to more fundamental and strategic matters by taking the UK exercise as a benchmarking model. In addition, the above four Korean exercises can draw good examples from the UK cases in relation to the utilisation type of the evaluation findings. Regarding evaluation methods, the Korean evaluation systems for the research councils and their member research institutes may adopt aspects of the UK experience, such as public consultation in the quinquennial review of research councils and the bibliometric analysis in the 2001 IAE.

## **8. 2. Limitations and Future Research**

As in other researches in the area of social sciences, there are some limitations in this thesis and further research need to be done in order to overcome them. The limitations of this research are of two types : one relates to the research methods adopted in this thesis, the other concerns the actual process of the application of the principal suggestions of the thesis to the Korean cases.

First of all, the limitation of the research methods will be briefly outlined. In this research, both a literature review and interviews were adopted as principal research methods. However, in fact, the literature review, which was conducted in a qualitative way, was the main research method, whereas the interviews with the related staff in the two countries functioned as supplementary methods. As a result, it seems that the two case studies and their meta-evaluations focused more on the formal structures and processes of the individual evaluation systems that were described in the official documents, including the related evaluation manuals and evaluation reports, than on actual structures and processes that might be inconsistent with the formal ones during evaluation exercises.

The evaluation exercise, in particular, at the level of policy and large-scale programmes, is considered as a typical social process both inside the pertinent institutions and among stakeholders. Therefore, there is an extra need to identify the actual structures and processes, if there is any possibility of this situation arising although it does not occur frequently. For this purpose, the interviews with the related evaluators and the researchers who were involved in the policy, programmes and institutions to be evaluated are considered as a useful research method.

For this reason, I also conducted in-depth interviews with related officials and staff in the UK and Korea face-to-face or by e-mail and telephone in individual cases, in order to identify these processes that are inconsistent with the formal ones when they occur in the UK or the Korean exercises. However, it seems that this effort was not sufficient for the full achievement of the initial intentions established at the research planning stage of my thesis. This shortcoming of the interviews resulted partly from the restrictions on the time schedule of the interviewees and in part from their attitudes of not willing to fully disclose the related information.

Next, there are limitations related to the fact that, when the principal suggestions of this thesis are directly applied to the future development of the Korean R&D evaluation systems, there will be some need to elaborate and specify these suggestions and make more detailed decisions at the operating level. This is because the suggestions for the applicability of aspects of the UK exercises to the Korean cases are discussed in a more or less broad and inclusive way.

The individual five R&D evaluation exercises adopted as the case studies in the UK and Korea include many evaluation components, even though this situation is somewhat inevitable in order to systematically and comprehensively understand the overall R&D evaluation systems of the two countries at national level. Therefore, it seems that each evaluation exercise in the UK and Korea has a broad and large enough research scope to be dealt with as a separate and independent research topic.

In consequence, further research need to be undertaken to overcome the above limitations of the thesis. Firstly, in order to grasp the actual structure and process during the evaluation exercise as well as the formal ones described in the official documents, there is

a need to slightly supplement some suggestions of the thesis by adopting more varied research methods, including a site visit to the institutions concerned and an extensive consultation with the related researchers and staff. This is because these are usually considered as more effective and suitable research methods for investigating the informal or unexpressed intentions of the evaluation exercises, which are not easy to understand from the review of formal documents as well as from interviews which last a relatively short time.

Secondly, in order to obtain more relevant and realistic information concerning the application of aspects of the UK case to the Korean exercise, there is a need to conduct detailed individual case studies in relation to some important UK R&D evaluation exercises, where appropriate. Examples might be the evaluation of funding for UK civil space activities and the quinquennial review of the six grant awarding research councils. These two UK evaluation exercise are expected to be able to provide the Korean cases with very useful references.

### **8.3. Contributions**

Despite the above limitations, this research will make two types of contribution, one from an academic perspective and the other from a practical one. From the academic perspective, it will contribute to the development of both meta-evaluation research in the public sector and comparative studies in the R&D evaluation system. Meanwhile, in practice, this research will assist in making R&D evaluation activities more effective and rational both in the UK and Korea. It also aims to make a contribution to cooperation between the UK and Korea in the area of science and technology policy in broader terms.

#### ***8.3.1. Academic Perspective : Development of Research in the Area of Meta-evaluation and R&D Evaluation***

The contributions from the academic perspective are of two types : one relates to the meta-evaluation research and the other concerns the comparative and regional studies in the area of S&T policy, in particular, R&D evaluation systems.



First of all, regarding the meta-evaluation research, it appears that this research has suggested a more systematic and comprehensive meta-evaluation model than existing research. It may also be the first case of the meta-evaluation conducted in relation to the overall R&D evaluation system at national level. The former matter relates to the theoretical development in the area of meta-evaluation research, whereas the latter is concerned with the specific and detailed case study in this field.

In order to develop the research framework, namely a meta-evaluation model, suitable for the research purposes of my thesis, I reviewed some existing research that deals directly or indirectly with this subject, covering such factors as components, standards, principles and the framework of meta-evaluation (Cook and Gruder, 1978 ; Stufflebeam, 1981 ; Larson and Berliner, 1983 ; The Joint Committee, 1994 ; AEA, 1995 ; Rogers, 1995). However, this research has had certain limitations which make it difficult to adapt directly to the design of a meta-evaluation model for my thesis. So, I have designed a new meta-evaluation model which had integrated and re-organised individual evaluation components discussed in existing research as well as proposing new components of meta-evaluation such as the evaluation paradigm made up of three sub components : an evaluation purpose, an evaluation object and evaluation types. It is evident that this research framework was able to deal with a very wide range of evaluation components from the evaluation paradigm to the evaluation utilisation of the individual evaluation system. In particular, my meta-evaluation model is expected to be a relevant reference for the development of theories discussed by Larson and Berliner (1983) and Rogers (1995), which directly focused on components of meta-evaluation and its framework.

Meanwhile, there have been many case studies conducted by adopting meta-evaluative approaches at the level of policy, programmes, projects and institutions (White et al., 1984 ; Bobe, 1991 ; Georghiou, 1995 ; Georghiou, 1999 ; Horvat, 1999 ; Hong, 2000 ; Yi and Knag, 2000b). Of these case studies, some have considered only a few components which have been simply derived from the basic concept of meta-evaluation, that is, the evaluation of evaluation, without any effort to design a research framework suitable for the characteristics of individual evaluation objects as well as the nature of the primary evaluation conducted. Therefore, it is apparent that the contribution of such case studies to the theoretical development of meta-evaluation is of limited significance.

Despite many case studies in this field, it seems that there have been few studies in relation to the evaluation of evaluation systems themselves at national level. The evaluation of evaluation systems also needs to be conducted, along with the evaluation of individual policy, programmes and institutions. This is because it can provide related researchers and policy makers with comprehensive information which they can use to develop the entire evaluation system itself, although that information is somewhat broad and too inclusive. For this reason, on the basis of a more comprehensive meta-evaluation model comprising four main components and 11 sub ones, this research dealt with the five evaluation systems making up the overall national R&D evaluation system of the UK and Korea. In this respect, my research is considered as one of the meaningful case studies in the area of meta-evaluation research, which mainly aimed to systematically understand evaluation systems themselves of certain countries.

Consequently, my thesis is expected to contribute to the theoretical development of meta-evaluation research by designing a new meta-evaluation model suitable for the evaluation of an evaluation system, in particular, a national R&D evaluation system, as well as by verifying its actual adaptability and generality through actually conducting the evaluation of five evaluation exercises in the UK and Korea.

Next, this research is expected to contribute to the accumulation of comparative and regional studies in the area of S&T policy, in particular, regarding the R&D evaluation systems of the UK and Korea. While, for Korea, this research will be one of the comparative studies between that country and other developed countries, in the UK, it will be also regarded as one of the East Asian regional studies.

As in other research fields, there is also much comparative research in relation to S&T policy between the UK and Korea, which is mainly conducted by Korean researchers. However, there are only a few concerning the national R&D evaluation systems of the two countries. This research adopted the five evaluation exercises of the two countries as case studies and compared them on the basis of the 11 evaluation components of the meta-evaluation model in order to identify which components are different and better in the UK and in Korea. Therefore, it is apparent that this research is considered as a well-timed comparative study to provide the advanced experiences of the UK to the related people and organisations responsible for R&D evaluation matters in Korea,

bearing in mind that the establishment of more rational R&D evaluation systems and their effective implementation are a currently emerging issue in the Korean S&T sector.

Meanwhile, for the UK, this research will contribute to the accumulation of regional studies of S&T affairs relating to countries whose relations with the UK have been recently strengthened and enhanced. It is apparent that Korea is one of the countries that undertake much collaborative work, in particular, in the area of S&T affairs. In particular, in order to give a general understanding of the Korean S&T affairs, this research has covered the S&T governance structure, the overview of S&T related ministries and some important statutes concerning S&T matters as well as the R&D evaluation system and its actual exercises. As a result, the research will provide people and organisations involved in SET in the UK with relevant knowledge in relation to the Korean S&T sector.

### ***8.3.2. Practical Perspective : Development of the R&D Evaluation Systems of the UK and Korea***

From the practical perspective, the findings of this research will be utilised for the development of the five sub evaluation systems making up the national R&D evaluation system in both the UK and Korea. In addition, it will contribute to mutual cooperation between the UK and Korea in the area of science and technology policy in broader terms.

First of all, even though the principal concern of this research is on the identification of the applicability of aspects of the UK exercise to the Korean case, it also contains a lot of information relevant to the development of the UK system. In particular, I have evaluated the five UK evaluation exercises in the SET sector in order to identify which components are the strengths and limitations in the individual UK R&D evaluation systems. This meta-evaluation shows that some UK evaluation systems, including the evaluation of the funding of UK civil space activities and the quinquennial review of the six research councils, are considered as good evaluation exercises, whereas certain evaluation exercise such as the cross-cutting review of science and research needs to be improved in some evaluation components. Consequently, it is evident that these findings of the meta-evaluations are expected to be good references for the future development of the UK R&D evaluation systems, even though they are discussed from the viewpoint of an

overseas researcher.

Next, as in the UK cases, this research will assist in making the Korean R&D evaluation systems and their exercises more rational and effective. As in the UK cases, I have identified some strengths and limitations of the individual Korean evaluation systems by conducting meta-evaluations. As a result, it seems that there is much need to improve the systems themselves and their actual exercises in most Korean R&D evaluation systems, excluding the evaluation system for the creative research initiatives (CRI) programme. Therefore, this research will provide related staff and institutions with essential information for the future development of the overall Korean R&D evaluation systems.

For this purpose, I have suggested that four out of the five Korean exercises, i.e. excluding the evaluation system of R&D project, can gain useful information from the UK cases. I also pointed out that some UK exercises can provide the Korean cases with useful references in relation to some evaluation components, in particular, an evaluation object, the utilisation types of the evaluation findings and evaluation methods.

In addition, some findings of this research are expected to be directly or indirectly utilised as basic knowledge to promote mutual cooperation between the UK and Korea in the area of S&T policy in broader terms. The case studies and their meta-evaluations in the UK and Korea include not a little general information in relation to such matters as governance system and its policy process in the S&T area. Therefore, it seems that when individual countries aim to promote cooperative works in the area of S&T affairs, they can obtain some basic information about their partners from this research.

## BIBLIOGRAPHY

- Alemi, Farrokh (1988), "Subjective and Objective Methods of Evaluating Social Programmes", *Evaluation Review*, 11(6), pp.765-774.
- Alkin, Marvin C., Daillak, Richard and White, Peter (1979), *Using Evaluations : Does Evaluation Make a Differences?*, Beverly Hills, California : Sage Publications.
- American Evaluation Association, Task Force on Guiding Principles for Evaluators (1995), "Guiding Principles for Evaluators", in William R. Shadish, Dianna L. Newman, Mary Ann Scheirer and Christopher Wye (eds.), *Guiding Principles for Evaluators (New Directions for Program Evaluation, No.66)*, San Francisco : Jossey-Bass, pp.19-26.
- Anderson, Scarvia and Ball, Stanley (1978), *The Profession and Practice of Programme Evaluation*, San Francisco : Jossey-Bass.
- Apthorpe, Raymond and Gasper, Desmond (1982), "Policy Evaluation and Meta-evaluation : The Case of Rural Co-operation", *World Development*, 10(8), pp.651-668.
- Arvidsson, Göran (1986), "Performance Evaluation", in F-X Kaufmann et al. (eds.), *Guidance, Control and Evaluation in the Public Sector*, Berlin : Walter de Gruyter.
- Barber, John M. (1999a), "Evaluation of Support by the UK Department of Trade and Industry for the Exploitation of Science, Technology and Innovation by UK Civil Industry", TESE (Technology, Economics, Statistics and Evaluation) Directorate, DTI [available at the TESE website (<http://www.dti.uk/tese/>)].
- Barber, John M. (1999b), "Creating an Anglo-Saxon Innovation Culture", in Susanne Bühner and Stefan Kuhlmann (eds.), *Evaluation of Science and Technology in the New Europe : Proceedings of an International Conference on 7 and 8 June 1999, Berlin, Karlsruhe, Germany : Fraunhofer ISI*, pp.33-44.
- Barber, John (2001), "Role of Appraisal and Evaluation in Policy-making", in Philip Shapira and Stefan Kuhlmann (eds.), *Proceedings from the 2000 US-EU Workshop on Learning from Science and Technology Policy Evaluation*, Bad Herrenalb, Germany : School of Public Policy, GeorgiaTech and Fraunhofer ISI, pp.9.24-9.33.

- Barkdoll, Gerald L. (1980), "Type III Evaluations : Consultation and Consensus", *Public Administration Review*, 40(2), pp.174-179.
- Barker, D. and Lloyd, P. (1997), "Evaluation of Scientific Research in the United Kingdom", *The Evaluation of Scientific Research : Selected Experiences*, OECD Report, pp.47-58.
- Barker, Katherine (1997), "The British Model : Evaluation by Professionals", in Michel Callon, Philippe Laredo and Mustar Philippe (eds.), *The Strategic Management of Research and Technology : Evaluation of Programmes*, Paris : Economica International, pp.55-68.
- BBSRC (Biotechnology and Biological Sciences Research Council, The UK) (1999a), *Institute Assessment Exercise 2001 : Guidelines to Institutes on the Procedure and Information Requirements for Submissions to the 2001 Institute Assessment Exercise (IAE)*.
- BBSRC (1999b), *Strategic Plan (1999-2004)*.
- BBSRC (2000), "Review of the Gastrointestinal Response to Food (GRF) Programme" (Internal Document), Assessment and Policy Unit.
- BBSRC (2001a), *Annual Report (2000-2001)*.
- BBSRC (2001b), "Evaluation of BBSRC Initiatives" (Internal Document), Assessment and Policy Unit.
- BBSRC (2001c), "BBSRC Research Grants (Revised September 2001)", <http://www.bbsrc.ac.uk/funding/research/handbook/booklet.htm> (3 October 2001).
- BBSRC (2002a), "About BBSRC", <http://www.bbsrc.ac.uk/about/> (21 February 2002 ).
- BBSRC (2002b), "Report of Meeting of the BBSRC Council : 12 February 2002", <http://www.bbsrc.ac.uk/news/council/> (26 February 2002 ).
- BBSRC (2002c), "The Science We Support", <http://www.bbsrc.ac.uk/science/> (1 March 2002).
- BBSRC (2002d), "BBSRC Publishes 4-yearly Institute Assessment Exercise (IAE) Result", [http://www.bbsrc.ac.uk/news/pressreleases/02\\_45\\_15\\_iaeweb.html](http://www.bbsrc.ac.uk/news/pressreleases/02_45_15_iaeweb.html) (5 May 2002).
- BBSRC (2002e), "Outcome of the BBSRC's 2001 Institute Assessment Exercise (IAE)", <http://www.bbsrc.ac.uk/about/centres/iae2001/web.html> (5 May 2002).
- BNSC (British National Space Centre) (2002), "UK Space Policy : Space Evaluation Report (Government Response)", <http://www.bnsc.gov.uk/> (3 March 2002).

- Bobe, B. (1991), "Trends in the Use of R&D Output Indicators in EC Programme Evaluation", *Sicentometrics*, 21(3), pp.263-282.
- Bozeman, Barry and Melkers, Julia (eds.) (1993), *Evaluating R&D Impacts : Methods and Practice*, Boston : Kluwer Academic Publishers.
- British Council (1999), *A Guide to the Organisation of Science and Technology (Gost) in Britain*, Manchester, UK.
- Bryson, John M. and Cullen, John W. (1984), "A Contingent Approach to Strategy and Tactics in Formative and Summative Evaluation", *Evaluation and Programme Planning*, 7(3), pp.277-289.
- Cabinet Office (1989), *R&D Assessment : A Guide for Customers and Managers of Research and Development*, London : Her Majesty's Stationery Office (HMSO).
- Cabinet Office (1999a), *The Modernising Government*, London : HMSO.
- Cabinet Office (1999b), *Professional Policy Making for the Twenty First Century*, London : HMSO.
- Cabinet Office (2000a), *Guidance on Agency and NDPB Quinquennial Review : How to Review Agencies and Non-Departmental Public Bodies (NDPBs) to Improve the Quality and Effectiveness of Public Services*, <http://www.cabinet-office.gov.uk/eeg/2000/review/contents.htm> (1 October 2001).
- Cabinet Office (2000b), *Writing It Up : Whitehall's Management of Cross-cutting Policies and Services*, London : HMSO.
- Cabinet Office (2000c), *Adding It Up*, London : HMSO.
- Cabinet Office (2002a), "Guide to the Centre of Government : Part II : Who Does What in the Treasury", <http://www.cabinet-office.gov.uk/roleofcentre/treasury.htm> (14 January 2002).
- Cameron, Hugh et al. (1996), *Evaluation of the Information Engineering Advanced Technology Programme : Summary Report*, Manchester : PREST, The University of Manchester.
- Campbell, Donald T. and Stanley, Julian C. (1966), *Experimental and Quasi-experimental Design*, Chicago : Rand McNally.
- Cha, Eui-hwan (1998), *Theory and Practice of Policy Evaluation : Approach and Strategy for Institution Evaluation* (in Korean), Seoul : Hanul Academy Publishing.
- Chelmsky, Eleanor (ed.) (1985), *Program Evaluation : Patterns and Directions*, Washington, D. C. : The American Society for Public Administration.

- Chelimsky, Eleanor (1987), "The Politics of Program Evaluation", in David S. Cordray, Howard S. Bloom and Richard J. Light (eds), *Evaluation Practice in Review (New Directions for Program Evaluation, No. 34)*, San Francisco : Jossey-Bass Publishers, pp.5-21.
- Chung, Sunyang and Grupp, Hariolf (1990), "R&D Policies in Germany and their Evaluation (I) : R&D Promotion Policies and Evaluation Approaches", STEPI, *Journal of Science and Technology Policy*, 2(1/2), pp.141-167.
- Cook, Thomas D. (1974), "The Potential and Limitations of Secondary Evaluation", in Michael W. Apple, Michael J. Subkoviak and Henry S. Lufner, Jr. (eds), *Educational Evaluation : Analysis and Responsibility*, Berkeley, California : McCutchan Publishing Corporation, pp.155-235.
- Cook, Thomas D. and Gruder, Charles L. (1978), "Metaevaluation Research", *Evaluation Quarterly*, 2(1), pp.5-51.
- Cooksy, Leslie J. (1999), "The Meta-Evaluand : The Evaluation of Project TEAMS", *American Journal of Evaluation*, 20(1), pp.123-136.
- Cunio, K. M. (1995), "UK Government Departments Experience of RT&D Programme Evaluation and Methodology", *Scientometrics*, 34(3), pp.363-374.
- Cunningham, Paul (ed.) (1998a), *Science and Technology in the United Kingdom (second edition)*, London : Catermill International.
- Cunningham, Paul (1998b), "The Organisation of UK Science and Technology", in Paul Cunningham (ed.), *Science and Technology in the United Kingdom (second edition)*, Bath : Cartermill International, pp.69-126.
- Cunningham, Paul (1998c), "The Higher Education System", in Paul Cunningham (ed.), *Science and Technology in the United Kingdom (second edition)*, Bath : Cartermill International, pp.159-204.
- Curran, Vennon R. (2000), "An Eclectic Model for Evaluating WEB-based Continuing Medical Education Courseware System", *Evaluation and the Health Professions*, 23(3), pp.318-347.
- Dale, Alison and Barker, Katherine (1994), "The Evaluation of EUREKA : A Pan-European Evaluation of a Pan-European Collaborative Technology Programme", *Research Evaluation*, 4(2), pp.66-74.
- De Bant, Jacques (1995), "Research and Innovation : Evaluation Problems and Procedures at Different Levels", *International Journal of Technology Management*, 10(4/5/6), pp.365-377.



- Dolowitz, David P. et al. (2000), *Policy Transfer and British Social Policy : Learning from the USA?*, Buckingham : Open University Press.
- Dolowitz, David P. (2000), "Policy Transfer : A New Framework of Policy Analysis", in David P. Dolowitz et al. (2000), *Policy Transfer and British Social Policy : Learning from the USA?*, Buckingham : Open University Press, pp.9-37.
- Dror, Yehezkel (1971), *Ventures in Policy Sciences*, New York : American Elsevier Publishing Co.
- DTI (The Department of Trade and Industry, The UK) (2000a), *Guidance on Preparing Evaluation Plans*, Internal Document : ES Central Evaluation Team.
- DTI (2000b), *Excellence and Opportunity : A Science and Innovation Policy for the 21st Century*, London : HMSO.
- DTI (2001a), *DTI Science and Innovation Strategy 2001*.
- DTI (2001b), *UK Research Councils : Research Institutes*, London : HMSO.
- DTI (2001c), *Evaluation of Funding for UK Civil Space Activity : An Overall Report and Reports by the Individual Funding Partners* (Assessment Paper No. 42).
- DTI/Assessment Unit (AU) (2001), *DTI Financial Support for UK Civil Space Activity (Chapter Two)*, in the *Evaluation of Funding for UK Civil Space Activities*, DTI.
- DTI (2002a) "Technology, Economics, Statistics and Evaluation (TESE) Directorate", <http://www.dti.gov.uk/tese/> (1 March 2002).
- DTI/OST (2001a), *Quinquennial Review of the Council for the Central Laboratory of the Research Councils : Report on the Stage One*.
- DTI/OST (2001b), *Quinquennial Review of the Grant Awarding Research Councils : Report on the Stage One*.
- DTI/OST (2001c), *Quinquennial Review of the Grant Awarding Research Councils : Stage Two Report*.
- DTI/OST (2002), *Quinquennial Review of the Council for the Central Laboratory of the Research Councils : Stage Two Report*.
- ECRC (Evaluation Committee for Research Councils) and the KISTEP (2001), *Evaluation Report of Research Councils in 2000* (in Korean), Seoul, Korea.
- European Commission (2000), *Good Practice Guidelines for the Management of the Evaluation Function*, Brussels : EU ([http://europa.eu.int/comm/budget/evaluation/en/goodpractice\\_en.pdf](http://europa.eu.int/comm/budget/evaluation/en/goodpractice_en.pdf)).
- Evaluation Research Society (USA), Standards Committee (1982), "Evaluation Research Society Standards for Program Evaluation", in Peter Rossi (ed.)

*Standards for Evaluation Practice (New Directions for Program Evaluation)*, San Francisco : Jossey-Bass.

Evered, David and Harnett, Sara (eds.) (1989), *The Evaluation of Scientific Research (Ciba Foundation Conference)*, Chichester : John Wiley and Sons.

Filstead, William J. (1979), "Qualitative Methods : A Needed Perspective in Evaluation Research", in Thomas D. Cook and Charles S. Reichardt (eds.), *Qualitative and Quantitative Methods in Evaluation Research*, Beverly Hills, California : Sage Publications.

Flanagan, Kieron and Keenan, Michael (1998), "Trends in UK Science Policy", in Paul Cunningham (ed), *Science and Technology in the United Kingdom (2nd edition)*, Bath : Cartermill International, pp.21-68.

Gallegos, Arnold (1994), "Meta-evaluation of School Evaluation Models", *Studies in Educational Evaluation*, 20, pp.41-54.

Geisler, Eliezer (1994), "Key Output Indicators in Performance Evaluation of Research and Development Organization", *Technological Forecasting and Social Change*, 47, pp.189-203.

Generics (2002), "Who We Are : Group Companies – Scientific Generics –", <http://www.generics.co.uk> (9 March 2002).

Georghiou, Luke and Cameron, Hugh (1987), *The ESPRIT Programme and the UK (Report to the Alvey Directorate, Department of Trade and Industry)*, PREST, The University of Manchester.

Georghiou, Luke (1989), "Organization of Evaluation", in David Everd and Sara Harnett (eds.), *The Evaluation of Scientific Research (Ciba Foundation Conference)*, Chichester : John Wiley & Sons, pp.16-31.

Georghiou, Luke and Meyer-Krahmer, Frieder (1992), "Evaluation of Socio-economic Effects of European Community R&D Programmes in the SPEAR Network", *Research Evaluation*, 2(10), pp.5-15.

Georghiou, Luke (1995), "Assessing the Framework Programmes : A Meta-evaluation", *Evaluation*, 1(2), pp.171-188.

Georghiou, Luke (1998), "Issues in the Evaluation of Innovation and Technology Policy", *Evaluation*, 4(1), pp.37-51.

Georghiou, Luke G. (1999), "Meta-evaluation : Evaluation of Evaluations", *Scientometrics*, 45(3), pp.523-530.

- Georghiou, Luke and Roessner, David (2000), "Evaluating Technology Programs : Tools and Methods", *Research Policy*, 20, pp.657-678.
- Gibbons, Michael and Georghiou, Luke (1987), *Evaluation of Research : A Selection of Current Practices*, OECD : Paris.
- Glaser, Edward and Backer, Thomas (1972), "Outline of Questions for Program Evaluators Utilizing the Clinical Approach", *Evaluation*, 1(1).
- Goldenberg, Edie N. (1983), "The Three Faces of Evaluation", *Journal of Policy Analysis and Management*, 2(4), pp.515-525.
- Grasso, Patrick G. (1999), "Meta-evaluation of an Evaluation of Reader Focused Writing for the Veterans Benefits Administration", *American Journal of Evaluation*, 20(2), pp.355-370.
- Greene, Jennifer C. (1992), "A Case Study of Evaluation Auditing as Metaevaluation", *Evaluation and Program Planning*, 15, pp.71-74.
- Guy, Ken and Georghiou, Luke et al. (1991), *Evaluation of the Alvey Programme for Advanced Information Technology*, London : HMSO.
- Guy, Ken and Arnold, Erik (1993), "UK Government Practices in Science and Technology Evaluation", *Research Evaluation*, 3(3), pp.176-186.
- Hager, Willi and Hasselhorn, Marcus (1998), "The Effectiveness of the Cognitive Training for Children from a Differential Perspective : A Meta-evaluation", *Learning and Instruction*, 8(5), pp.411-438.
- Ham, Christopher and Hill, Michael (1993), *The Policy Process in the Modern Capitalist State (2nd edition)*, London : Harvester Wheatsheaf.
- Hill, Michael (ed.) (1998), *The Policy Process : A Reader (2nd edition)*, London : Prentice Hall/Harvester Wheatsheaf.
- Hills, Philip (1995), "PREST's Experience of Evaluation", *Scientometrics*, 34(3), 401-414.
- Hills, Philip and Dale, Alison (1995), "Research and Technology Evaluation in the United Kingdom", *Research Evaluation*, 5(1), pp.35-44.
- Hills, Philip (2001), "European Study Report : United Kingdom", in Paul Cunningham et al.(eds.) *Measuring and Ensuring Excellence in Government Science and Technology : International Practices (France, Germany, Sweden and the United Kingdom)*, Manchester : PREST, The University of Manchester, pp.63-76.
- HM Treasury (Her Majesty's Treasury, the UK) (1983), *Financial Management in Government Departments* (Cmnd 9058), London : HMSO.

- HM Treasury (1988), *Policy Evaluation : A Guide for Managers*, London : HMSO.
- HM Treasury (1991), *Economic Appraisal in Central Government : A Technical Guide for Government Departments*, London : HMSO.
- HM Treasury (1997), *The Green Book : Appraisal and Evaluation in Central Government (Treasury Guidance)*, London : HMSO.
- HM Treasury (1999), *Creating Knowledge Creating Wealth : Realising the Economic Potential of Public Sector Research Establishments (John Baker Report)*, <http://www.hm-treasury.gov.uk/docs/1999/baker.htm> (24 April 2001).
- HM Treasury (2000a), *2000 Spending Review : Cross-cutting Study of Science Research Funding Analysis, Arguments and Proposals*, [http://www.hm-treasury.gov.uk/sr2000/associated/science\\_231100.pdf](http://www.hm-treasury.gov.uk/sr2000/associated/science_231100.pdf) (9 October 2001).
- HM Treasury (2000b), "Introduction to the Public Spending Framework", [http://www.hm-treasury.gov.uk/Documents/Public\\_Spending\\_and\\_Services/pss\\_intro.frame.cfm](http://www.hm-treasury.gov.uk/Documents/Public_Spending_and_Services/pss_intro.frame.cfm) (14 January 2002).
- HM Treasury (2000c), *The 2000 Spending Review Public Service Agreements White Paper (2001-2004)*, London : HMSO.
- HM Treasury (2001a), "HM Treasury Science and Innovation Strategy 2001", <http://www.hm-treasury.gov.uk/docs/2001/sciencestrat.html> (9 October 2001).
- HM Treasury (2001b), "Set Out Priorities for 2002 Spending Review", <http://www.hm-treasury.gov.uk/press/2001/p72-01.html> (18 September 2001).
- HM Treasury, DfES, DTI and OST (2002), *Cross-Cutting Review of Science and Research : Final Report*, [http://www.ost.gov.uk/policy/advice/science\\_cross\\_cutter.pdf](http://www.ost.gov.uk/policy/advice/science_cross_cutter.pdf) (22 October 2002).
- HM Treasury (2003), "Recent Publications : Green Book (2003 edition)", [http://www.hm-treasury.gov.uk/recent\\_pubs](http://www.hm-treasury.gov.uk/recent_pubs) (16 January 2003).
- Hogwood, C. and Gunn, L.A. (1984), *Policy Analysis in the Real World*, Oxford : Oxford University Press.
- Hollister Jr., Robinson G. et al. (1979), "Linking Process and Impact Analysis", in Thomas Cook and Charles Reichardt (eds.), *Qualitative and Quantitative Methods In Evaluation Research*, Beverly Hills, California : Sage Publications, pp.140-158.
- Hong, Heung-deug and Kang, Keun-bok (1998), "The Evaluation of Large-scale and National R&D Programmes : Biotech 2000 Programme Case (in Korean)", *Journal of Korea Technology Innovation Society*, 1(1), pp.125-146.

- Hong, Heung-deug (2000), "Meta-evaluation of National Large-scale R&D Programmes : A Comparative Analysis of Evaluation Systems of 6 National R&D Programmes", Ph.D. Thesis, PREST, The University of Manchester.
- Hoogerwerf, Andreis (1992), "Policy Evaluation and Government in the Netherlands : Meta Evaluation Research as One of the Solutions", in John Mayne et al. (eds.), *Advancing Public Policy Evaluation : Learning from International Experiences*, Amsterdam : North-Holland, pp.215-227.
- Horvat, M. (1999), "Meta-evaluation : Evaluation of Evaluations Some Points for Discussion", *Scientometrics*, 45(3), pp.53-542.
- Hudson, Joe, Mayne, John and Thomlison, R. (eds.) (1992), *Action-oriented Evaluation in Organization : Canadian Practices*, Toronto, Ontario : Wall and Emerson.
- Hwang, Young-soo et. al. (2000), *A Comparative Analysis of the Government-supported R&D Programmes in six Countries and the Improvement of the Evaluation System* (in Korean), Seoul : MOST and STEPI.
- Jenkins, Bill (1998), "Policy Analysis : Models and Approaches", in Michael Hill (ed.), *The Policy Process : A Reader (2nd edition)*, London : Prentice Hall/Harvester Wheatsheaf, pp.30-38.
- Jones, Bill et al. (2001), *Politics UK (Edition 4)*, Harlow : Pearson Education Limited.
- KEDI (The Korean Educational Development Institute) (2002), "The Educational Statistics System : Educational Statistics Annual" (in Korean), <http://std.kedi.re.kr/html/frame1.html> (1 May 2002).
- KISTEP (The Korea Institute of Science and Technology Evaluation and Planning) (2002a), "Homepage", <http://www.kistep.re.kr/main00/english/> (10 April 2002).
- KISTEP (2002b), "Introduction to the KISTEP's Services" (in Korean), <http://www.kistep.re.kr/> (10 April 2002).
- Kim, Gyeo-soo, Kim, Jae-young and Lee, Min-hyung (1992), "The Evaluation System of the Strategic R&D Programmes" (in Korean), *Science and Technology Policy*, 4(2), pp.65-92, Seoul : STEPI.
- Kim, Jae-young (1999), *A Study on the Evaluation System of National R&D Projects Sponsored by Ministry of Science and Technology* (in Korean), Seoul : KISTEP
- Kim, Myung-soo (1993), *Public Policy Evaluation* (in Korean), Seoul : Bakyoung-Sa.
- Kim, Myung-soo et al. (1995), *Public Auditing*, Seoul : Daeyoung Moonhwa International.

- KOCI (The Korea Research Council for Industrial Science and Technology) (2002a), "Homepage", <http://www.koci.re.kr/english/> (1 May 2002).
- KOCI (2002b), *Manual for the Evaluation of Member Research Institutes in 2001* (in Korean), Seoul, Korea.
- KOCI (2002c), *Evaluation Report of Member Research Institutes in 2001* (in Korean), Seoul, Korea
- KORP (The Korea Research Council of Public Science and Technology) (2002), "Homepage", [http://www.korp.re.kr/e\\_index.html/](http://www.korp.re.kr/e_index.html/) (1 May 2002).
- Kostoff, Ronald N. (1994), "Quantitative/Qualitative Federal Research Impact Evaluation Practices", *Technical Forecasting and Social Change*, 45, pp.189-205.
- KRCF (The Korea Research Council of Fundamental Science and Technology) (2002), "Homepage", <http://www.kocf.re.kr/eng/> (1 May 2002).
- Larson, Richard and Berliner, Leni (1983), "On Evaluating Evaluations", *Policy Sciences*, 16(2), pp.147-163.
- Lasswell, Harold (1951), "The Policy Orientation", in Daniel Lerner and Harold Lasswell (eds.) *The Policy Sciences : Recent Developments in Scope and Method*, Stanford : Stanford University Press, pp.3-15.
- Lee, Jae-ok et al. (1993), *The Analysis of Performance in Large-scale R&D Project : TDX and TICOM Case* (in Korean), Seoul : STEPI.
- Lee, Jay-young (2002a), "Strategic Management of National R&D Programmes in Korea (II)", Seoul : KISTEP.
- Lee, Jin-joo et al. (1994a), *A Study on the Development and Implementation of R&D Evaluation Model* (in Korean), Daejeon, Korea.
- Lee, Jin-joo and Suh, Kun-soo (1996), *A New Model for the Assessment of Policy Reform*, (in Korean) Seoul : Namam Publishing House.
- Lee, Jung-won, (2000), *A Framework for R&D Evaluation System and its Application* (in Korean), Seoul : STEPI.
- Lee, Kong-rae (1999), *Characteristics of Korea's National Innovation System*, Seoul : Science and Technology Policy Institute (STEPI).
- Lee, Linda J. and Sampson, John F. (1990), "A Practical Approach to Program Evaluation", *Evaluation and Program Planning*, 13, pp.157-164.
- Lee, Min-hyung (2001b), *A New Approach in Organisational Performance Measurement System of Government-supported Research Institutes* (in Korean), Seoul : STEPI.

- Lee, Min-hyung (2002b), "Comprehensive Performance Evaluation System on Government R&D Institutes (GRIs) : Current Status and New Approach for Development", presented in *The 1st Joint Workshop between the Korea Research Council for Industrial Science and Technology (KOI) and the National Institute of Advanced Industrial Science and Technology (AIST) of Japan*, 19-20 November 2002, Jeju, Korea.
- Lee, Mu-shin et al. (1994b), "Effective R&D Evaluation System in the National R&D Programmes (in Korean)", *The Korean Journal of Management*, 23(special), pp.193-227.
- Lee, Mu-shin et al. (1995), "The Comprehensive Evaluation and Implementation of Information and Telecommunication R&D Programme" (in Korean), *The Korean Journal of Information Science*, 13(3), pp.102-111.
- Lee, Mu-shin et al. (1996), "Evaluation of National R&D Projects in Korea", *Research Policy*, 25, pp.805-818.
- Levitan, Sar A. and Warzburg, Gregory (1979), *Evaluating Federal Social Programs*, Kalamazoo, Michigan : The W. E. Upjohn Institute for Employment Research.
- Luukkonen, Terttu and Ståhle, Bertel (1990), "Quality Evaluation in the Management of Basic and Applied Research", *Research Policy*, 19, pp.357-368.
- Mackay, Keith (1992), "The Use of Evaluation in the Budget Process", *Australian Journal of Public Administration*, 51(4), pp.436-439.
- McGraw, Sarah et al. (1989), "Methods in Programme Evaluation : The Process Evaluation System of the Pawtucket Heart Programme", *Evaluation Review*, 13(5), pp.459-483.
- Meyer-Krahmer, Frieder (1995), "Technology Policy Evaluation in Germany", *International Journal of Technology Management*, 10(4/5/6/), pp.601-621.
- MIC (The Ministry of Information and Communication, The Republic of Korea) et al. (1998), *The Analysis of Performance in the Information and Communication R&D Programmes* (in Korean), Seoul, Korea.
- Miller, Roger (1992), "The Influence of Primary Task on R&D Laboratory Evaluation : A Comparative Bibliometry Analysis", *R&D Management*, 22(1), pp.3-20.
- MOCIE (The Ministry of Commerce, Industry and Energy, The Republic of Korea) (1992), *The Evaluation of the Industrial Generic R&D Programme (1)* (in Korean), Seoul, Korea.

MOE (The Ministry of Education and Human Resources Development, The Republic of Korea) (2002a), "Educational Reform Initiatives : Brain Korea 21", [http://www.moe.go.kr/eng\\_26/](http://www.moe.go.kr/eng_26/) (10 April 2002).

MoR (The Ministry of Reconstruction, The UK) (1918), *Report of the Machinery of Government Committee (Haldane Report)*, London : HMSO.

MOST (The Ministry of Science and Technology, The Republic of Korea) (1997), *The Evaluation of Major Policies in Korean Government : Biotechnology R&D Programme* (in Korean), Seoul, Korea.

MOST (1998), "The Special Act for Scientific and Technological Innovation" (in Korean), Seoul, Korea.

MOST (2000a), *Vision 2025 : Korea's Long-term Plan for Science and Technology Development*, Seoul, Korea.

MOST (2000b), "Master Plan for Analysis, Evaluation and Pre-Budget Review of National R&D Programmes in 2001" (in Korean), Seoul, Korea.

MOST (2001a), "The Framework Act on Science and Technology", "Enforcement Decree" and "Ministerial Ordinance" (in Korean), Seoul, Korea.

MOST (2001b), "The Presidential Decree on Planning, Management and Evaluation of National R&D Programmes" (in Korean), Seoul, Korea.

MOST (2001c), "The Regulations on Operation of the Evaluation Committee for National R&D Programmes" (in Korean), Seoul, Korea.

MOST (2001d), "The Technology Development Promotion Act", (in Korean), Seoul, Korea.

MOST (2002a), *Science and Technology Annual 2001* (in Korean), Seoul, Korea.

MOST (2002b), "Introduction to the MOST", [http://www.most.go.kr/index\\_e.html](http://www.most.go.kr/index_e.html) (10 April 2002).

MOST (2002c), "Science and Technology Policy", [http://www.most.go.kr/index\\_e.html](http://www.most.go.kr/index_e.html) (10 April 2002).

MOST (2002d), "Statistics of R&D in Science and Technology", [http://www.most.go.kr/index\\_e.html](http://www.most.go.kr/index_e.html) (10 April 2002).

MOST (2002e), "Government-supported Research Institutes", [http://www.most.go.kr/index\\_e.html](http://www.most.go.kr/index_e.html) (10 April 2002).

MOST, (2002f) "The Regulations on Management of National R&D Programme by the MOST" (in Korean), Seoul, Korea.



- MOST (2002g), "The Regulations on Evaluation for National R&D Projects by the MOST" (in Korean), Seoul, Korea.
- MOST (2002h), "Introduction to the National R&D Programme of the MOST" (in Korean), <http://www.most.go.kr/index.html> (1 October 2002).
- MOST and KISTEP (2000a), "Operating Plan of Interim Output Evaluation for the Creative Research Initiatives (CRI) Selected in 1997" (in Korean), Seoul, Korea.
- MOST and KISTEP (2000b), "Report of Interim Output Evaluation for the CRI Selected in 1997" (in Korean), Seoul, Korea.
- MOST (2002i), "Science and Technology Policy : International S&T Cooperation", [http://www.most.go.kr/index\\_e.html](http://www.most.go.kr/index_e.html) (15 December 2002).
- MPB (The Ministry of Planning and Budget, The Republic of Korea) (2002a), "Homepage", [http://www.mpb.go.kr/index\\_eng.html](http://www.mpb.go.kr/index_eng.html) (15 April 2002).
- Nachmias, David (1979), *Public Policy Evaluation : Approaches and Methods*, New York : St. Martin's.
- Nagel, Stuart (1987), "Evaluating Public Policy Evaluation", *Policy Studies Journal*, 16(2), pp.219-232.
- NSTC (The National Science and Technology Council, The Republic of Korea), (2001a), *Report of Analysis and Evaluation of National R&D Programmes in 2001* (in Korean), Seoul, Korea.
- NSTC (2002a), "Homepage" (in Korean), <http://www/nstc.go.kr/> (15 April 2002).
- Nyden, Michael (1992), "Evaluation of R&D in Sweden", in John Mayne et al. (eds.), *Advancing Public Policy Evaluation : Learning from International Experience*, Amsterdam : North-Holland, pp.159-171.
- OECD (1983), *Report of the Workshop on the Evaluation of Government Measures for the Stimulation of Innovation*, Paris, 6-7 June, 1983.
- OECD (1996), *Reviews of National Science and Technology Policy : Republic of Korea*, Paris, France.
- OECD (1997), *Report of the Conference on the Policy Evaluation in Innovation and Technology*, Paris, 26-27 June, 1997.
- OECD (2000), "OECD Science, Technology and Industry Outlook 2000 : Korea", <http://www.oecd.org/oecd/pages/document/> (15 October 2001).
- OGPC (The Office for Government Policy Co-ordination, The Republic of Korea) (2001a), *Manual for the Evaluation of the Research Councils in 2000* (in Korean), Seoul, Korea.

- OGPC and PEC (2001a), *Introduction to the Evaluation System of Government Performance* (in Korean), Seoul, Korea.
- OGPC and PEC (2001b), *Guidance for Government Performance Evaluation in 2001* (in Korean), Seoul, Korea.
- OGPC and PEC (2002a), *Comprehensive Report of Government Performance Evaluation in 2001* (in Korean), Seoul, Korea.
- OGPC and PEC (2002b), *Individual Report of Government Performance Evaluation in 2001 : The Case of the MOST* (in Korean), Seoul, Korea.
- OGPC and PEC (2002c), "Implementation Plan of Corrective Measures in Government Performance Evaluation in 2001" (in Korean), Seoul, Korea.
- OPM of Japan (The Office of the Prime Minister, Japan) (1997), *National Guideline on the Method of Evaluation for Government R&D*, Tokyo, Japan.
- OPM (The Office of the Prime Minister, The Republic of Korea) (2001a), "The Framework Act on Government Performance Evaluation" and "Enforcement Decree" (in Korean), Seoul, Korea.
- OPM (2001b), "The Regulations on Operation of the Policy Evaluation Committee" (in Korean), Seoul, Korea.
- OPM (2001c) "The Act on Establishment, Operation and Promotion of the Government-supported Research Institutes (GRIs)" and "Enforcement Decree" (in Korean), Seoul, Korea.
- OPM (2002a), "Homepage : Organisations in the Office of the Prime Minister", <http://www.opm.go.kr/home/english/> (10 April 2002).
- Orata, Pedro T. (1940), "Evaluating Evaluation", *Journal of Educational Research*, 33(9), pp.641-661.
- Ormala, Erkki (1989) "Nordic Experience of the Evaluation of Technical Research and Development", *Research Policy*, 18, pp.333-342.
- OST (The Office of Science and Technology, The UK) (1993), *Realising our Potential : A Strategy for Science, Engineering and Technology*, London : HMSO.
- OST (1997), *The Use of Scientific Advice in Policy Making : A Note by the Chief Scientific Advisor*, London : HMSO.
- OST (2000a), *Science and Engineering Base Group Management Plan 2000/2001*, London : HMSO.
- OST (2000b), *Guidelines 2000 : Scientific Advice and Policy Making*, <http://www.dti.gov.uk/ost/aboutost/guidelines.htm> (23 May 2001).

- OST (2001a), *Scientific Advice and Policy Making : Implementation of Guidelines 2000*.
- OST (2001b), "Announcement of Cross-cutting Review of Science and Research", <http://www.dti.gov.uk/ost/whatsnew/middlew.htm> (17 September 2001).
- OST (2001c), "SET Statistics 2001", <http://www.gov.uk/setstats/> (1 October 2001).
- OST (2002), "About OST", <http://www.dti.gov.uk/ost/aboutost/sebg.htm> (20 February 2002).
- Palumbo, Dennis and Nachmias, David (1983), "The Precondition for Successful Evaluation : Is There an Ideal Paradigm", *Policy Sciences*, 16, pp.67-79.
- Park, Jung-taek (1996), "Conceptual Characteristics and Major Components of Policy Paradigm", *Social Science Review*, Taejon University, The Social Science Research Institute, pp.39-57.
- Park, Young-ill (1995), "Planning, Implementation and Evaluation of the Large-scale R&D Programmes : Focused on the Frontier R&D Programme (in Korean)", Ph.D. Thesis, Daejon : The Korea Advanced Institute of Science and Technology (KAIST).
- Parsons, Wayne (1997), *Public Policy : An Introduction to the Theory and Practices of Policy Analysis*, Cheltenham : Edward Elgar Publishing Limited.
- Patton, Michael Quinn (1990), *Qualitative Evaluation and Research Methods*, Newbury Park, California : Sage Publications.
- PEC (The Policy Evaluation Committee, The Republic of Korea) (2002a), "Home Page", <http://pec.pec.go.kr/english/> (16 May 2002).
- PEC (2002b), "Introduction to the PEC and its Activities" (in Korean), <http://www.pec.go.kr> (1 June 2002).
- Praestgaard, E. (1999), "Meta-evaluation : Evaluation of Evaluations", *Scientometrics*, 45(3), pp.531-532.
- RAE (2001), "Results : View by Institutions", <http://www.rae.ac.uk/results/byinist.asp> (18 December 2001).
- Rho, Hwa-joon et al. (1996), *Development of Evaluative Components and its Weight for the Comprehensive Evaluation of Research Institute* (in Korean), Seoul : STEPI.
- Rich, Robert F. (1977), "Uses of Social Science Information by Federal Bureaucrats : Knowledge for Action versus Knowledge for Understanding", in Carol Weiss (ed.), *Using Social Research in Public Policymaking*, Lexington : Lexington Books.

- Rogers, Patricia J. and Hough, Gary (1995), "Improving the Effectiveness of Evaluations : Making the Link to Organizational Theory", *Evaluation and Program Planning*, 18(4), pp.321-332.
- Rogers, Patricia (1995), "Meta-evaluation : Evaluating the Evaluation and Evaluation System", paper in *Health Evaluation Conference*, Sydney, Australia.
- Rossi, Peter H. and Freeman, Howard E. (1993), *Evaluation : A Systematic Approach*, Newbury Park, California : Sage Publications.
- Rothschild, Victor (1971), "The Organisation and Management of Government R&D" (Rothschild Report), in *A Framework for Government Research and Development (Cmnd 4814)*, London : HMSO.
- Rutman, Leonard (1980), *Planning Useful Evaluations : Evaluability Assessment*, Newbury Park, California : Sage Publications.
- Sanders, James R. (1995), "Standards and Principles", in William R. Shadish, Dianna L. Newman, Mary Ann Scheirer and Christopher Wye (eds.), *Guiding Principles for Evaluators (New Directions for Program Evaluation, No.66)*, San Francisco : Jossey-Bass, pp.47-52.
- Sanders, James R. (1999), "Meta-evaluation of the Effectiveness of Comprehensive, Case Management Interventions : Evidence from the National Evaluation of the Comprehensive Child Development Program", *American Journal of Evaluation*, 20(3), pp.577-582.
- Schwandt, Thomas and Halpern, Edward (1988), *Linking Auditing and Metaevaluation : Enhancing Quality in Applied Research*, London : Sage Publications.
- Schneider, Anne Larason (1986), "The Evolution of a Policy Orientation for Evaluation Research : A Guide to Practice", *Public Administration Review*, 46(4), pp.356-363.
- Scheirer, Mary Ann (1994), "Designing and Using Process Evaluation", in Joseph S. Wholey, Harry P. Hatry and Kathryn E. Newcomer (eds.), *Handbook of Practical Programme Evaluation*, San Francisco : Jossey-Bass Publishers, pp.40-68.
- Scriven, Michael (1969), "An Introduction to Meta-evaluation", *Educational Product Report*, No.2, pp.36-38.
- Scriven, Michael (1991), *Evaluation Thesaurus (fourth edition)*, Newbury Park, California : Sage Publications.
- Shadish Jr, William R., Cook, T. D. and Leviton, L. C. (1991), *Foundations of Program Evaluation : Theory and Practice*, London : Sage Publications.

- Shattock, Michael L. (1991), "Higher Education and the Research Councils", in Robin Nicholson, Catherine M. Cunningham and Philip Gummatt (eds.), *Science and Technology in the United Kingdom*, Essex : Longman, pp.44-68.
- Smith, N. and Hauer, W. (1990), "The Applicability of Selected Evaluation Models to Evolving Investigative Designs", *Studies in Educational Evaluation*, 16, pp.489-500.
- Smith, Nick L. (1990), "Citations on the Use of Investigative Case Studies in Meta-evaluation", *Evaluation and Program Planning*, 13, pp.373-378.
- STEPI (Science and Technology Research Institute, The Republic of Korea) (2002), "Introduction to the STEPI and its Research Activities", (in Korean), <http://www.stepi.re.kr> (15 August 2002).
- Strube, Michael J. and Hartmann, Donald P. (1983), "Meta-analysis : Techniques, Applications and Functions", *Journal of Consulting and Clinical Psychology*, 51(1), pp.14-27.
- Stufflebeam, Daniel (1981), "Metaevaluation : Concept, Standard, and Uses", in Ronald A. Berk, (ed.), *Educational Evaluation Methodology*, The Johnes Hopkins University Press, pp.146-163.
- Suchman, Edward A. (1967), *Evaluation Research : Principle and Practice in Public Service and Social Action Programmes*, New York : Russell Sage Foundation.
- Suchman, Edward A. (1972), "Action for What? A Critique of Evaluative Research", in Carol H. Weiss (ed.), *Evaluating Action Programs*, Boston : Allyn & Bacon.
- The Joint Committee on Standards for Educational Evaluation (1988), *The Personnel Evaluation Standards : How to Assess Systems for Evaluating Educators*, Newbury Park, California : Sage Publications.
- The Joint Committee on Standards for Educational Evaluation (1994), *The Program Evaluation Standards : How to Assess Evaluations of Educational Programs (2nd Edition)*, London : Sage Publications.
- Technopolis Group (2001), *Overarching Report (Chapter One)*, in the *Evaluation of Funding for UK Civil Space Activities*, London : DTI.
- Technopolis (2002), "Homepage", <http://www.technopolis.co.uk/index.htm> (5 March 2002)
- United Kingdom Evaluation Society (UKES) (2001), "Changing the Green Book : HM Treasury's Guide to Appraisal and Evaluation in Central Government", [http://www.evaluation.org.uk/ukes\\_new/UKES/Evaluator\\_Article1.htm](http://www.evaluation.org.uk/ukes_new/UKES/Evaluator_Article1.htm) (17 Dec. 2001).

- Uusikylä, Petri and Virtanen, Petri (2000), "Meta-Evaluation as a Tool for Learning", *Evaluation*, 6(1), pp.50-65.
- Vedung, Evert (1997), *Public Policy and Program Evaluation*, London : Transaction Publishers.
- Waller, Mike (1992), "Evaluating Policy Advice", *Australian Journal of Public Administration*, 51(4), pp.440-446.
- Weinberg, Alvin M. (1989), "Criteria for Evaluation : A Generation Later", in David Everd and Sara Harnett (eds.) *The Evaluation of Scientific Research (Ciba Foundation Conference)*, Chichester : John Wiley and Sons, pp.3-15.
- Weiss, Carol H. (1972), *Evaluation Research : Methods of Assessing Program Effectiveness*, Englewood Cliffs, N. J. : Prentice-Hall.
- White, Lawrence T. et al. (1984), "Energy Conservation Research of California's Utilities : A Meta-evaluation", *Evaluation Review*, 8(2), pp.167-186.
- Wholey, Joseph S., Hatry, Harry P. and Newcomer, Kathryn E. (eds.) (1994), *Handbook of Practical Program Evaluation*, San Francisco : Jossey-Bass Publishers.
- Yang, Hee-seung (2002), "The Coordination of National R&D Programmes In Korea : Evaluation and Pre-Budget Review", Seoul : KISTEP.
- Yi, Chan-goo (1997), "Meta-evaluation on the Information and Communication R&D Programme in Korea (in Korean)", *Korean Public Administration Quarterly*, 9(4), pp.651-667.
- Yi, Chan-goo and Kang, Keun-bok (1998), "Directions for Improvements in the Evaluation System of IC R&D Program in Korea (in Korean)", *Korean Policy Studies Review*, 7(1), pp.97-121.
- Yi, Chan-goo and Kang, Keun-bok (2000a), "A Design of the Meta-evaluation Model", in *Canadian Evaluation Society's 20th Annual Conference*, 14-17 May 2000, Montreal, Canada.
- Yi, Chan-goo and Kang, Keun-bok (2000b), "Developments of the Evaluation System of Government-supported Research Institutes in Korean Science and Technology", *Research Evaluation*, 9(3), pp.158-170.
- Yi, Chan-goo (2002), *The Management System of the UK Research Councils and their Research Institutes : The Case of the BBSRC and the IAH* (in Korean), Seoul : STEPI.

# LIST OF INTERVIEWEES

## 1. The United Kingdom

Case Study	Organisation	Name
<b>Cross-cutting Review (2000)</b>	Enterprise Team HM Treasury	Andrew Morgan
	Technology and Innovation Team HM Treasury	Helen Watson
<b>Quinquennial Review (2001)</b>	Research Councils Directorate, SEBG OST	Gavin Costigan
	Research Councils Directorate, SEBG OST	Angelica Gundona
<b>IAE of the BBSRC (2001)</b>	Assessment and Policy Unit BBSRC	Mari Williams
	Assessment and Policy Unit BBSRC	Carol Miller
<b>UK Civil Space Activity (2001)</b>	Technology, Economics, Statistics and Evaluation Directorate (TESE) DTI	John Barber
	Assessment Unit, TESE DTI	Derek Albone
	Assessment Unit, TESE DTI	Nigel Gibbons
	TESE/DTI & UK Evaluation Society (former)	Philip Hills
<b>GRF Programme of the BBSRC (2000)</b>	Assessment and Policy Unit BBSRC	Mari Williams
	Assessment and Policy Unit BBSRC	Carol Miller
	Institute for Animal Health BBSRC	Geoff Oldham

## 2. The Republic of Korea

Case Study	Organisation	Name
<b>Evaluation of Major Policy Issues of the MOST (2001)</b>	Office of Director-General I for Policy Analysis and Evaluation OGPC	Sang-yoon Chang
<b>Evaluation of National R&amp;D Programmes (2001)</b>	Division of National Research and Development Evaluation KISTEP	Jang-jae Lee
<b>Evaluation of Five Research Councils (2001)</b>	Office of Director-General I for Policy Analysis and Evaluation OGPC	Sang-yoon Chang
	Division of National Research and Development Evaluation KISTEP	Jang-jae Lee
	Evaluation and Management Team KOCI	Mi-jeong Jung
<b>Evaluation of Member Research Institutes Of the KOCI (2002)</b>	Evaluation and Management Team KOCI	Mi-jeong Jung
<b>Evaluation of the Creative Research Initiatives of the MOST (2000)</b>	Division of National R&D Planning and Management KISTEP	Bu-jong Geel