

Comparative Analysis of Research and Development Evaluation on Government Ministries and Agencies*

Keita YAMAZAKI**

(Assistant Auditor, Agriculture Audit Division, Board of Audit)

1. Foreword

1.1 Purpose of Research

Recently, voices have been raised about the importance of evaluating the country's administration system, which also holds true for the research and development sector. In August 1997, the National Guidelines on the Method of Evaluation for Government Research and Development (hereafter referred to as the "National Guidelines on Evaluation") were laid down by the Prime Minister. Since then, government ministries and agencies began to put into practice their approach to the evaluation of research and development projects, but tackling R&D evaluation began only recently, and the method adopted varies with the organizations.

The purpose of this study is to characterize the R&D evaluation of respective government ministries and agencies and further probe the stance of the administration toward the evaluation of national R&D projects. The method chosen is mainly by comparative analysis of evaluation results released on the Internet web site, specifically with respect to current developments in evaluation of state R&D schemes.

1.2 Processes Leading to the Formulation of the National Guidelines on Evaluation

This paper provides a brief summary of the background behind the introduction of the National Guidelines on Evaluation, which triggered the activation of R&D project evaluation by ministries and agencies.

Institution of the Science and Technology Basic Law (November 1995)

The Law was enacted to push ahead with policies and measures related to science and technology in a comprehensive, planned and aggressive manner, in an attempt to build up Japan into a "nation of science and technology."

Formulation of the Science and Technology Master Plan (July 1996)

This Plan was approved by the Cabinet in July 1996 after the Government shouldered legal obligation to establish the Science and Technology Master Plan under Article 9.1 of the Science and Technology Basic Law.

Formulation of the National Guidelines on Evaluation (August 1997)

The Master Plan, stressing "the rigid enforcement of evaluation" as one of its key items, prescribed the introduction of the National Guidelines on Evaluation, which led to approval of the Guidelines by the Prime Minister.

1.3 Defining the Concept of Policy Evaluation

The evaluation of research and development resides within the broad framework of policy evaluation. There can be diverse views and ideas on policy evaluation and it is difficult to define it precisely. Besides the term "policy evaluation," terms such as "program evaluation" and "project evaluation" are often used. Therefore, this study

* This paper is a summary of the thesis for the master's degree that the author presented when dispatched to the Postgraduate Course of Policy Science Research of Saitama University for domestic research as an administrative officer of the National Personnel Authority in FY1999.

** Born in 1970; joined the Board of Audit in 1994 and assigned to Second Education Audit Division and Personnel Division before assuming current position.

starts off by defining the difference between “policy,” “program,” and “project.”

According to Kiyoshi Yamaya, “policies,” “programs,” and “projects” stand for three levels. First, “policy,” refers to “the representation of an idea by sector toward which an entire organization should proceed and is often described as “plans.” Next, “the means to achieve the policy goal” is “the program,” and “the means to achieve “the program goal” is “the project.” Yamaya argues that the “chain of policy to program and project” is thereby built up into a single system ¹⁾.

Evaluation is applied to each of these levels for the purpose of determining if the goal of the policy, program, or project is appropriate or if the costs incurred are reasonable, etc., specifically to improve the quality of the policy, program, or project. Then, specific evaluators, timings, techniques, etc. are established for each level as the means to conduct these evaluations.

1.4 Overview of National R&D evaluation

National research and development is carried out based on the specific missions of ministries and agencies. It is classified into two; one is research and development carried out by competent national research institutes and corporations having a special status, and the other is research and development implemented by the internal subdivisions of ministries and agencies under R&D system and large-scale R&D projects, etc.

National research institutes are assigned various R&D projects as either routine ²⁾ or special ³⁾ research schemes. Likewise, the internal government divisions perform specific assignments under the R&D system, and large-scale R&D projects.

Such research and development evaluation involves project evaluation and institute evaluation, where “project evaluation” refers to the evaluation of R&D projects carried out by national research institutes as routine and special research, and the evaluation of R&D assignments carried out by the internal government divisions under the R&D system and large-scale R&D projects. On the other hand, institute evaluation refers to the evaluation of research institutes in terms of policy and management, specifically in the fields of organization and personnel administration, the selection of R&D fields and projects, the improvement of facilities, information infrastructure and research support system, the promotion of exchange with external organizations (including joint research), etc. Figure 1 shows these interrelations.

In term of the three levels of “policy,” “program,” and “project” mentioned above, the majority of objects of research and development belong to the “project” level, since most of them cover concrete R&D schemes. As to “institute evaluation,” since the presence of a research institutes itself is considered a “program” to implement research and development schemes, “institute evaluation” can be regarded as “program” level evaluation.

Further, the evaluation of national research and development schemes is often consigned to an external evaluation committee comprising outside evaluators elected by national research institutes or the internal subdivisions of the government. Occasionally, evaluation is performed by researchers of the institutes evaluated, internal subdivision staff, etc.

As to the timing of evaluations, project evaluation is conducted before, during, and after implementation of projects; specifically, “pre-evaluation,” “interim-evaluation,” and “post-evaluation.” On the other hand, institute evaluations are generally conducted every 3 to 5 years.

There are diverse methods for the evaluation of research and development projects. Basically, these can be classified into qualitative and quantitative evaluation techniques, where the former covers peer reviews, case studies, and survey studies, and the latter covers checklists, evaluation marking and mathematical planning techniques, as well as “portfolios” which are “system theory” methods, financial and econometric techniques, and even philology-metric techniques.

Of these techniques, national research institutes often use the peer review for qualitative evaluation, which

1) Kiyoshi Yamaya “Theory of Policy Evaluation and Associated Developments” (P.11), published by Koyo Shobo, 1997.

2) Ordinary research: focused on relatively basic fields supporting the basis of all research activities financed by manpower-based research funds (per-head allocation according to fixed unit rate to cover expenses on routine research)

3) Special research: conducted on urgent projects to be completed within a specified timeframe and plans to meet social and administrative demands.

refers to the performance of R&D evaluation by other researchers or specialists who engage in the same fields of research. Exceptionally, evaluation of large-scale R&D projects called “mega-science” is joined by non-specialist scholars, etc. Peer reviews are often conducted because they are simple and less costly, but involve drawbacks such as the evaluation tends to be based on a subjective viewpoint.

For quantitative evaluations, gradation rating systems are often used. This refers to a method where evaluators mark achievements based on the predetermined rating standards of 3 to 5 gradations. This method is easier and affords more objective “transparent” evaluation, except that it, too, is partly based on the evaluator’s subjective view.

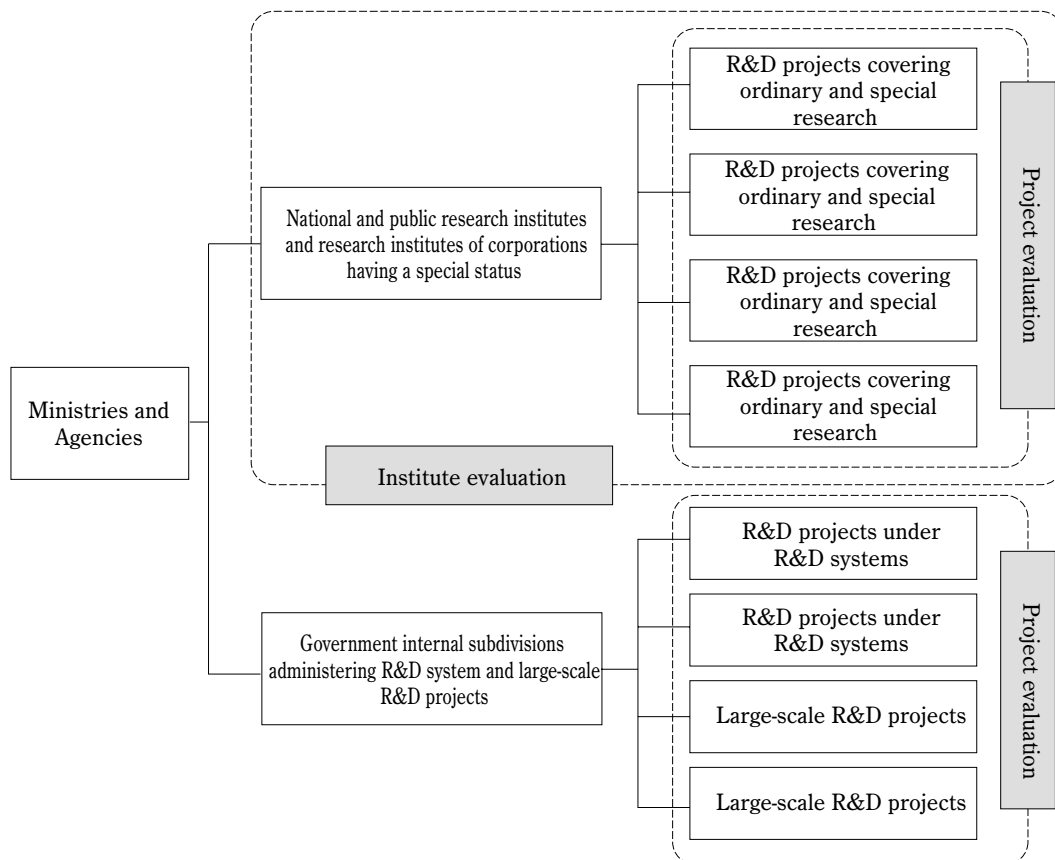


Figure 1 Objects and Schemes for National R&D Evaluation

1.5 Objects and Viewpoints of Analysis in this Study

This study analyzes the project evaluation and institute evaluation of national research institutes and the evaluation of projects carried out by internal subdivisions under the R&D system, including large-scale R&D projects.

However, since it is difficult to cover all of these, this study limits the scope of analysis to “evaluation information” relative to the results of evaluations released on the Internet Web sites of relevant governmental organizations and those of national research institutes, directing specific attention to “Internet disclosure.” The review of information was limited to those carried on the Internet Web sites for the five months from June 16 through November 15, 1999. Therefore, the phrase “at present” roughly refers to the end of 1999.

Then, with respect to viewpoints, the National Guidelines on Evaluation claim that attention should be directed to the following points of evaluation:

- Statement of evaluation standards and processes
- Introduction of “external evaluation”

- Enforcement of “open evaluation”
- Proper use of evaluation results, such as “Reflection of evaluation on effective allocation of R&D resources”

Based on these four factors, this study extracts 1) evaluation guidelines and procedures, 2) external evaluation, 3) evaluation results and 4) effective use of evaluation results as the key points for analysis, and analyzes specific items relating to the evaluation of research and development of applicable ministries and agencies.

2. Comparison of Governmental Organizations in Evaluation of Research and Development

2.1 R&D Evaluation by National Research Institutes

(1) Formulation and disclosure of evaluation guidelines and procedures

In response to the request for a “statement of evaluation standards and processes” under the National Guidelines on Evaluation, ministries and agencies kicked off formulation of evaluation guidelines and procedures. Here, the evaluation guidelines refer to the internal general guidelines of these organizations that lay down basic policies on their research and development programs, as compiled on ministry and agency levels. The evaluation procedures refer to the R&D systems administered by competent ministries and agencies and to concrete measures for evaluation individually laid down by national research institutes under the ministries and agencies.

First, the paper reviews formulation and disclosure of the state level evaluation guidelines. Among the nine governmental organizations that laid down evaluation guidelines were the Ministry of Posts and Telecommunications (Basic Plan for Research and Development of Information Communication Technology, April 1997), the Ministry of Agriculture, Forestry and Fisheries (Guidelines on Evaluation of National Research Institutes and Research Schemes, July 1997), the Ministry of International Trade and Industry (Guidelines on Technological Evaluation, August 1997), the Science and Technology Agency (Promotion of R&D Evaluation, September 1997), the Ministry of Education (Proposed Standards on Evaluation of Academic Research, December 1997), the Ministry of Health and Welfare (Guidelines on Enforcement of Evaluation relating to Welfare Science Research, January 1998), the Environment Agency (Basic Guidelines on Research Evaluation, February 1998), the Ministry of Transport (Evaluation Guidelines on Research and Development, February 1998) and the Defense Agency (Evaluation Guidelines ;Name of the Guidelines is uncertain, April 1998). These governmental organizations are privileged to have liberal budget allocations to cover science and technology research, and rank among the top ten ministries and agencies in terms of expenditure on science and technology (as of FY 1998). Of these nine ministries and agencies, there were four that released information on the Internet, including the Ministry of Posts and Telecommunications, the Ministry of Agriculture, Forestry and Fisheries, the Ministry of International Trade and Industry, the Science and Technology Agency, and the Ministry of Education.

With respect to the formulation and public disclosure of the evaluation guidelines of the national research institutes, the author finds that among 94 institutes, there were 36 institutes that laid down evaluation guidelines (including those in the process of formulation). Among the research institutes that have formulated guidelines, there were only 8 that released them on the Internet (five under the Science and Technology Agency, two under the Ministry of Transport, and one under the Fire Defense Agency of the Ministry of Home Affairs). On the other hand, there were many organizations for which it was impossible to determine the presence or absence of formulation. More likely, such organizations have not established specific evaluation procedures, but have been involved in evaluations under their general evaluation guidelines. The author was unable to find details on the facts of formulation. In this instance, the author would like to point out that the “National Guidelines on Evaluation” provide for the need to establish transparent, objective standards of evaluation in order to afford outsiders access to the details of evaluation. Nevertheless, at present, few evaluation guidelines laid down by the national research institute have been disclosed to date.

(2) Evaluation of R&D projects

i) Definition of “external evaluation”

There are various ideas about external evaluation. For example, according to Atsunobu Ichikawa, “internal evaluation” refers to the case where evaluation criteria are laid down by the decision-making organization, while

“external evaluation” refers to the case where evaluation criteria are established by outside organizations. But Ichikawa claims that evaluation cannot be classified as “internal” or “external” depending only on whether evaluators are outside or internal people.⁴⁾ He claims that evaluation should not be defined as “external” only because evaluators are outside people, but instead, it should be defined as “external” when the evaluation criteria established by outsiders is applied to the evaluation. On the other hand, the New Administrative System Research Association states “policy evaluation can be classified as “internal” or “external” depending on who performs the evaluation. In the case that the evaluation policy is planned and executed by the administration, it is defined as an “internal evaluation,” but where it is planned and executed by organizations outside the administration, it is defined as an “external evaluation.” Internal evaluation includes evaluations such as those performed by outside scholars, survey institutes, etc. on consignment to the administration. Evaluations conducted by the Administrative Inspection Bureau may be called “quasi-internal evaluations.” Outside evaluators can include the Diet, which has control over the administration, the Board of Audit, non-administrative policy advisory bodies, and civic groups.⁵⁾ Where the administration asks outsiders for evaluation, such a case is not defined as an external evaluation merely because outside evaluators were employed.

However, the National Guidelines on Evaluation states in the chapter, Introduction of External Evaluation, that “it is necessary to elect neutral third parties (belonging neither to the evaluating nor evaluated bodies) as outside evaluators in order to assure objective, equitable, and reliable evaluation.” The Guidelines defines evaluation as “external” on condition that the evaluators are selected from outside sectors. Such a case, however, is applicable to internal but not to external evaluations, according to Ichikawa and the New Administrative System Research Association. More strictly, the phrase “introduction of external evaluation” under the National Guidelines of Evaluation should be construed as meaning the “introduction of external evaluators.”

This study reviews the current progress in the introduction of outside evaluators invited under the National Guidelines on Evaluation.

ii) Introduction of external evaluation

Analysis was directed at 16 national research institutes, including the National Research Institute of Police Science of the National Police Agency; the National Research Institute for Metals, the National Institute of Radiological Science, the National Research Institute for Earth Science and Disaster Prevention and the National Research Institute of Inorganic Substances of the Science and Technology Agency; the National Research Institute of Brewing of the Ministry of Finance; the National Institute of Livestock and Agrobiological Science and the Tohoku National Fisheries Research Institute of the Ministry of Agriculture, Forestry and Fisheries; the Ship Research Institute, Port and Airport Research Institute, the National Traffic Safety and Environment Laboratory and the Meteorological Research Institute of the Ministry of Transport; the National Institute of Industrial Safety of the Ministry of Labor; the Public Works Research Institute and the Geographical Survey Institute of the Ministry of Construction; and the National Research Institute of Fire and Disaster of the Ministry of Home Affairs. [Introduction of outside evaluators]

Out of the 16 research institutes, 14 organizations practiced evaluation employing outside evaluators. In two institutes under the Ministry of Agriculture, Forestry and Fisheries, evaluation was conducted by internal staff. One institute (National Research Institute for Metals) employed internal as well as external evaluators at the Science and Technology Agency, plus one institute (Public Works Research Institute) at the Ministry of Construction.

4) Atsunobu Ichikawa "Research Evaluation" (P.25) (No. 46 of Series of Lectures on Theory of Research Evaluation by Research Institute of Science and Technology Policy, 1996.

Modification to the original version by the author.

Original: "Standard of Evaluation" refers to a norm which represents a yardstick for the best option. Internal evaluation refers to cases where the evaluation standards are established by a decision making body, while external evaluation refers to a case where they are established by external body (same as the original version in following paragraphs).

5) "Techniques of Policy Evaluation" compiled by the New Administrative System Research Association, released in the morning edition of Nihon Keizai Shimbun (May 10, 1999).

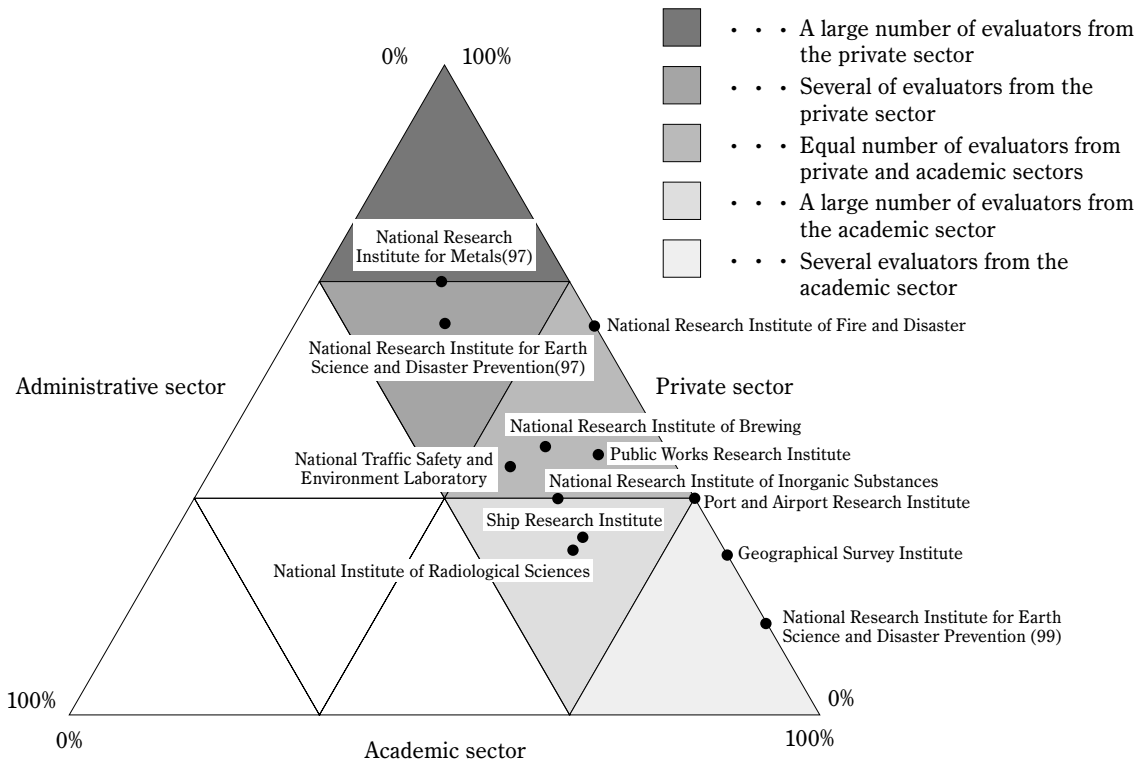


Figure 2 Composition of Outside Evaluators by Origin

[Composition of outside evaluators]

In respect of foreign evaluators, no positive introduction has been by governmental organizations, except that the Science and Technology Agency introduced them for two institutes (National Institute of Radiological Sciences and National Research Institute for Earth Science and Disaster Prevention) and the Ministry of Construction, for one institute (Geographical Survey Institute).

This paper classifies Japanese outside evaluators by origin of private, academic, and government sectors. Figure 2 breaks down Japanese outside evaluators by origin (private, academic, and administrative sectors) with respect to 11 research institutes that made public their position titles, and plots their proportions in a triangle graph. The graph shows a higher proportion of evaluators from the private sector toward the top, a higher proportion of evaluators from the academic sector towards the lower right and a higher proportion of evaluators from administrative sector towards the lower left.

As shown in the figure, evaluators from the private and academic sectors share roughly equal proportions, with those from academe perhaps exceeding those from the private sector by a narrow margin. A further breakdown of evaluators by origin in the private sector shows that more evaluators come from public service corporations including corporate juridical persons and juridical foundations. Of the 39 outside evaluators from the private sector, those from public service corporations total 20 evaluators (51.2%), including 11 (28.2%) from corporate juridical persons and 9 (23.0%) from juridical foundations.

Public service corporations are nonprofit foundations established under the civil law to serve public benefits and interests, e.g. in the sectors of rituals, religion, charity, academy, art, culture, etc. However, some of them are closely linked with governmental organizations, including those operating as state-affiliated organizations. Often, these organizations provide outlets for retired senior government officials, who employ them as directors, etc. Although corporate bodies operate under the civil law, some public service corporations are established with fundamental property donated by the Federation of Economic Organizations and its affiliates at the request of the competent administration. In other words, new public service corporations are established under the leadership of competent government agencies. Among public service corporations, there are some nonprofit organizations that

operate as economic groups which are not considered quasi-governmental organizations. In this sense, it would be unreasonable to conclude that all these public corporations are quasi-governmental organizations. However, the author assumes that a research institute is likely to select evaluators for the research institute from public corporations that have a connection with the institute, that is from quasi-governmental organization.

Based on what may be an extreme assumption, the author analyzed the proportion of evaluators, with public service corporations counted as quasi-governmental organization, or even as administrative organizations (Figure 3). In light of this, the proportion of outside evaluators weighing more on the private sector in Figure 2, changed to weigh more on the administrative sector as shown in Figure 3. In this instance, if a given government organization has close connections with a public service corporation and currently administers some research institute, evaluation by a member of such a public corporation should be characterized more as “internal.”

In fact, how objective and neutral are outside evaluators concerning the research institutes to be evaluated? For example, a person who acts as an outside evaluator at the time of an evaluation may be a retired officer of the government ministry or agency administering the research institute for which he is now serving as an evaluator. Also, there may be cases where a person from an academy is selected to be an evaluator of a research institute having close interests with that academy. There can be endless possibilities in this respect. The purpose of the introduction of outside evaluators is to assure objective, equitable and reliable evaluations, as stated in the “National Guidelines on Evaluation.” Should a person close to internal sources be selected as an outside evaluator, it could jeopardize such an objective evaluation. As mentioned by Ichikawa and the New Administrative Research Association, a given evaluation turns into an internal evaluation at the moment an evaluated organization selects its evaluator by itself. Specifically, it is no longer worthy of being termed an external evaluation in the true sense of the word.

In this sense, considerable attention should be given to the selection of personnel as outside evaluators. In reality, people have no alternative but to trust that outside evaluators are neutral in connections with the research institutes they serve. I believe that national research institutes should be paying particular attention to this issue. However, there is currently no means for checking that such evaluators undertake objective, equitable and reliable evaluations. Consequently, as of now, one realistic and feasible means for ensuring objective evaluations would be

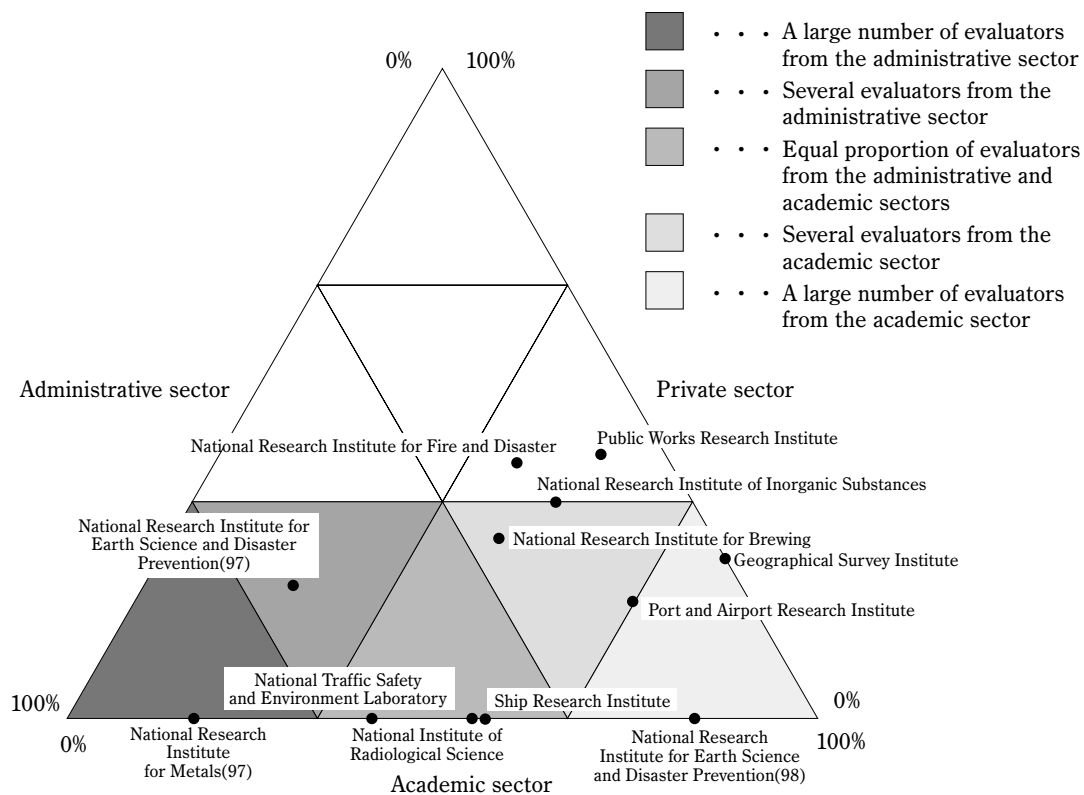


Figure 3 Composition of Outside Evaluators by Origin, defining Public Service Corporations as Part of the Administration

for national research institutes to take a positive stance to substantiate the neutrality of outside evaluators and establish evaluation guidelines and procedures that ensure the objective and neutral positions of evaluators. Such guidelines and procedures should then be disclosed to the public, possibly accompanied by the personal history of the outside evaluators concerned.

iii) Disclosure and contents of evaluation

[Extent of disclosure]

In this section, the paper analyzes how much information has been disclosed by each national research institute concerning the results of evaluations. In order to determine the extent of disclosure, evaluation results of these institutes disclosed on Internet Web sites were printed out on A4-sized paper and the number of pages output was used as the indicator of volume.

Figure 4 shows the extent of disclosure on the results of project evaluations by national research institutes, where the National Institute of Radiological Science (Science and Technology Agency) topped the list of disclosure with 63.8 pages, followed by the National Research Institute for Metals with 31.4 pages (FY 1998). Other institutes' disclosure averaged some five pages, although this does not provide an accurate yardstick for the extent of disclosure because it is in proportion to the number of projects evaluated. Accordingly, the author next attempts to analyze the extent of disclosure per project evaluated (Figure 5).

The National Research Institute for Earth Science and Disaster Prevention topped the list of disclosure with 7.7 pages (FY 1997), followed by the National Institute of Radiology Science with 2.3 pages and the National Research

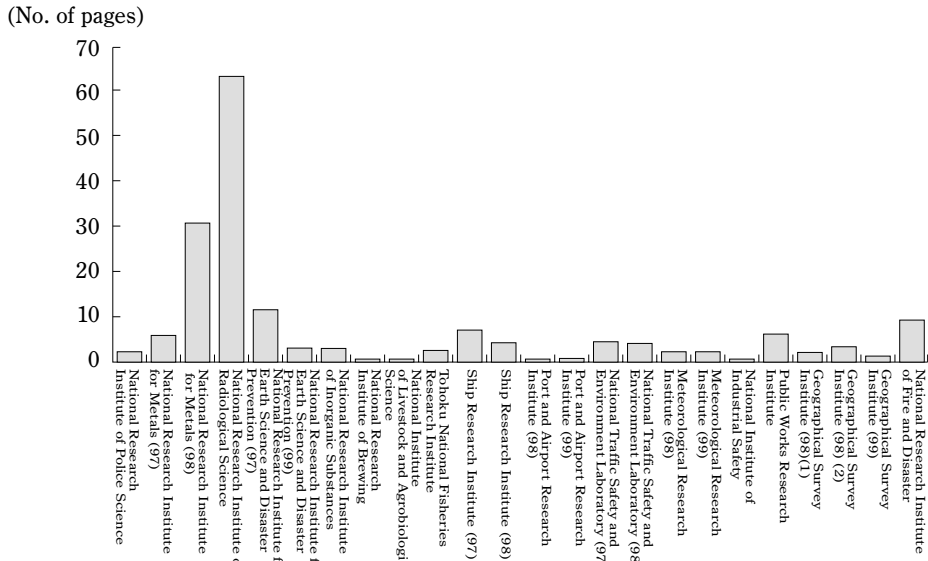


Figure 4 Volume of Internet Disclosure on Project Evaluation Results

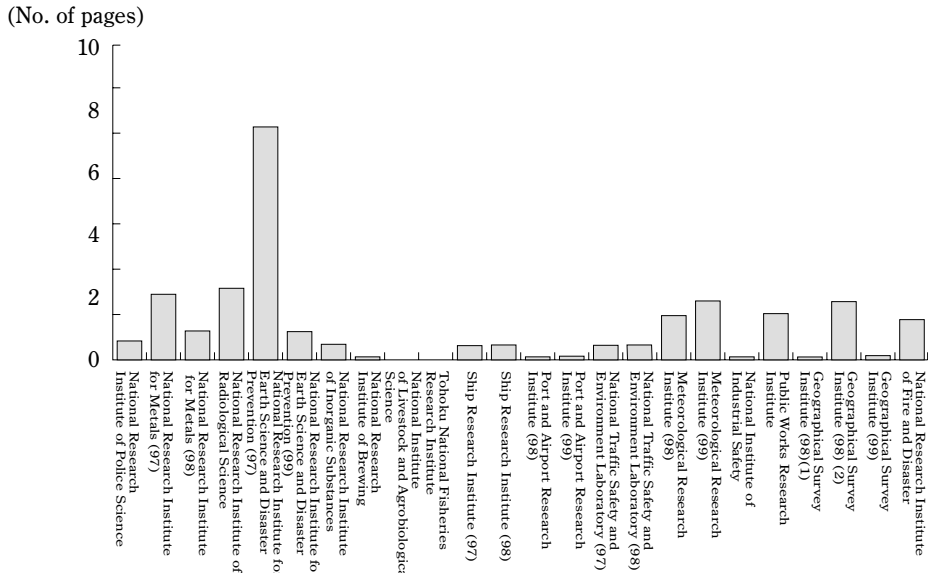


Figure 5 Extent of Disclosure of Evaluation Results per Project

Institute for Metals with 2.2 pages (FY 1977). Disclosure per project averaged 1.1 pages. In other words, most institutes came up with a disclosure of about one page.

Thus, common trend was that most national research institutes did not go beyond disclosure that comprised only brief summaries.

[Trends in presenting evaluation results]

In this section, the paper analyzes the method of presenting the disclosed evaluation results. Evaluation results cover results of item-wise evaluation and comprehensive evaluation. Some research institutes state both item-wise and comprehensive evaluation results, while others selectively use either method. Evaluation results are presented as descriptions or gradation ratings. If evaluation results are presented in gradation rating (as of a rating in an “ABC” order), it is easier for readers to understand evaluation reports. Furthermore, if this is supported by descriptive rating, it allows readers to garner more details.

On the other hand, if evaluation results are presented solely in descriptive ratings, this can obscure the meaning of such evaluation reports and readers may find it difficult to understand a report that often demands expertise or a professional background.

Investigation revealed that most research institutes presented evaluation reports using descriptive ratings, except for a few organizations that presented them in gradation rating of three to five scales—specifically those by the National Institute of Radiological Science, the National Research Institute for Metals, and two other institutes.

[Disclosure rates on pre-, interim- and post-evaluations]

Of the evaluation results disclosed, what are the proportions of disclosure with respect to pre-evaluation, interim-evaluation, and post-evaluation? Current analysis indicates high proportions of pre- and interim- evaluations, with post-evaluation showing a lower proportion. A high proportion of pre-evaluation is due perhaps to the concerns for the timing of budget requests. Similarly, a high proportion of interim-evaluations are due to their effects on the timing of budget requests to cover expenditure during the latter half of a project period.

[Contents of evaluation]

In this section, the paper presents the results of analysis on the contents of evaluation. It is often said that evaluation by ministries and agencies rarely generates negative results. Is this really true? The author analyzed this issue based on disclosed evaluation results. The scope of analysis was limited to research institutes that presented results in a relatively concise and comprehensible manner.

Here, evaluation results are rated in gradations of “high,” “ordinary,” and “low,” and each research institute determined the proportions of these ratings. Some institutes used the rating of “fair on the whole,” but the author classified this within the “ordinary” rating. Figure 6 plots results in a triangle graph.

As shown in the graph, the results of evaluations by the administration show an intensive distribution of results in the “high rating” area. As expected, no negative evaluation results were seen. The area of “fair on the whole,” or the ordinary rating, shows a dense distribution of research institutes under the Ministry of Transport. Particularly, research institutes under the Ministry of Transport were rated “fair on the whole.” As shown in the legend of the figure, research institutes rated “fair” were marked []. Few, if any, institutes were rated “low.” The National Institute of Livestock and Agrobiological Science under the Ministry of Agriculture, Forestry and Fisheries were rated “slightly unsatisfactory” against some research schemes.⁶⁾ Interestingly, the National Institute of Livestock and Agrobiological Science, whose evaluation was conducted by an internal evaluator, was rated as “low,” while no other institutes evaluated by outside evaluators were rated as “low.”

iv) Utilization of evaluation results

The author surveyed how far each research institute was prepared for the effective utilization of their evaluation results. Those national research institutes that presented policies and measures to answer these results of evaluation include seven organizations of the National Research Institute for Earth Science and Disaster Prevention (FY 1997) of the Science and Technology Agency, the National Research Institute of Brewing of the

6) National Institute of Livestock and Agrobiological Sciences rated evaluation results in four gradations of "satisfactory", "relatively satisfactory", "slightly unsatisfactory", and "unsatisfactory". This paper reclassified two ratings of "satisfactory" and "relatively satisfactory" in the original into "high" rating and "slightly unsatisfactory" and "unsatisfactory" in the original into "low" rating.

Ministry of Finance, the Tohoku National Fisheries Research Institute of the Ministry of Agriculture, Forestry and Fisheries, the Port and Airport Research Institute, the National Traffic Safety and the Environment Laboratory of the Ministry of Transport, the Public Works Research Institute of the Ministry of Construction, and the National Research Institute of Fire and Disaster of the Ministry of Home Affairs. Most of the institutes' countermeasures evaluated centered on the "Amendment and reassessment of research schemes," as taken by five institutes. Others include "Change of project name" (by two institutes), and "Changes in budget allocation" (by two institutes). They came up with policies and measures to follow the recommendations, but described them in vague and inconclusive expressions on the whole. Other institutes did not even get as far as presenting countermeasures for recommendation, perhaps in view of the fact that only a short time had passed since the evaluation.

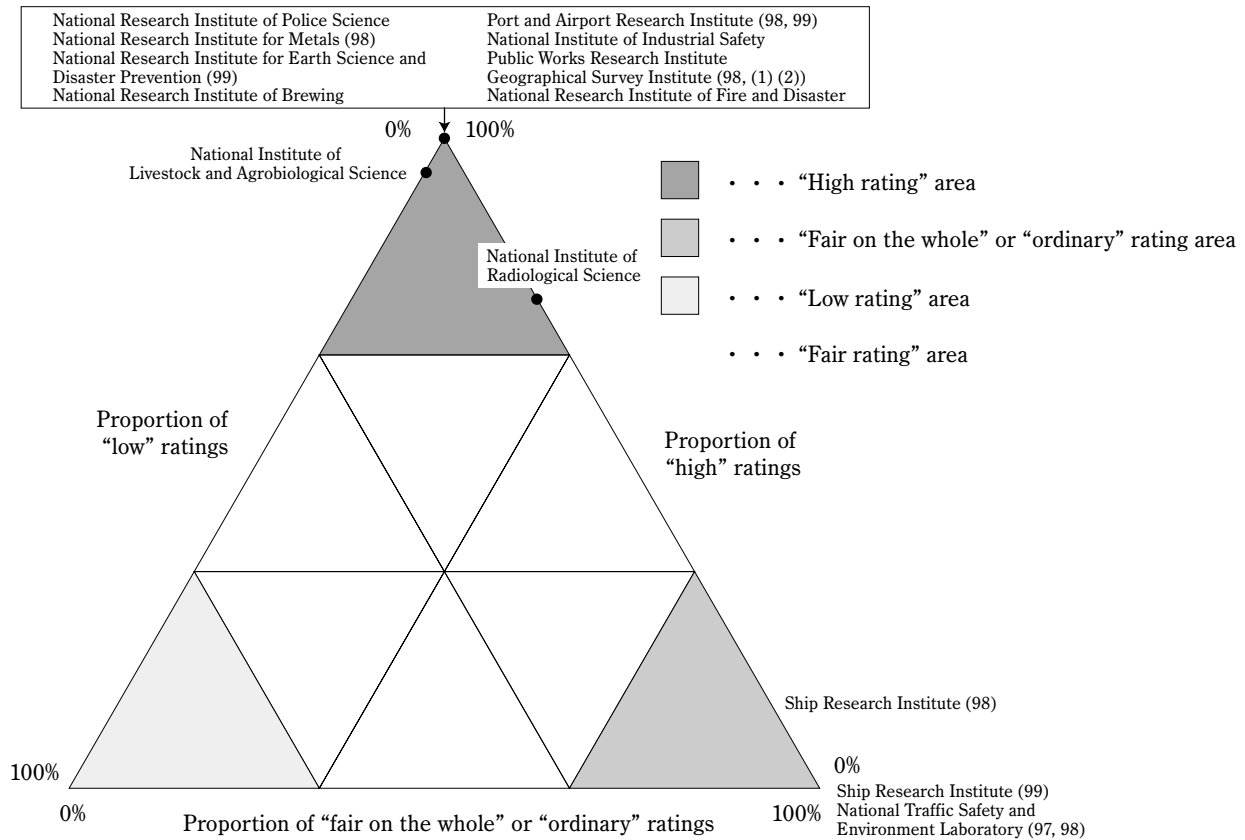


Figure 6 Breakdown of Evaluation Results into Three Ratings

(3) Implementation of institute evaluation

i) Introduction of external evaluation

Analysis was directed towards 31 national research institutes, including the National Aerospace Laboratory of Japan, the National Research Institute for Earth Science and Disaster Prevention, the National Research Institute of Inorganic Substances and the National Institute of Science and Technology Policy of the Science and Technology Agency; the Okazaki National Research Institute (Institute for Molecular Science) and the Science and Technology Information Center of the Ministry of Education; the Research Institute of National Rehabilitation Center for the Disabled and National Children's Hospital (National Children's Medical Research Center) of the Ministry of Health and Welfare; the National Institute of Agrobiological Sciences, the National Institute for Agro-Environmental Sciences, the National Institute of Livestock and Agrobiological Sciences, the National Institute of Fruit Tree Science, Hokkaido National Agricultural Experiment Station, the Hokuriku National Agricultural Experiment Station, the Shikoku National Agricultural Experiment Station, the National Institute of Animal Health, Forestry and Forest Products Research Institutes, the Hokkaido National Fisheries Research Institute, the Tohoku National Fisheries Research Institute, the National Research Institute of Fisheries and Environment of Inland Sea, the Seikai National Fisheries Research Institute, the Japan Sea National Fisheries Research Institute, the National

revealed that three out of 16 (18.8%) research institutes introduced them for project evaluations, whereas 10 out of 31 (32.3%) research institutes introduced them for institute evaluations. Thus they were accepted more for institute evaluations.

As mentioned, moves are underway to introduce foreign evaluators on the part of some research institutes, but this does not prevail in a majority of research institutes. The introduction of foreign evaluators can enhance the international reputations of those research institutes that accept them. On the other hand, these institutes sustain an extra workload and expenses, including the translation of reference evaluation materials into foreign languages for foreign evaluators, traveling expenses, etc. It therefore costs them more money and time. Accordingly, it would not be wise to unreservedly accept foreign evaluators without justifiable reasons. Therefore, upon the introduction of foreign evaluators, it is necessary to take diverse factors into consideration, centering on the means of achieving the objectives of the evaluation.

The paper now reviews the composition of Japanese evaluators by origin (whether from the private, academic, or administrative sector). Figure 7 indicates the dominant distribution of evaluators from the private sector with respect to the Science and Technology Agency and the Ministry of International Trade and Industry, and those from the administrative sector with respect to the Ministry of Agriculture, Forestry and Fisheries. The proportion of evaluators from public service corporations compared to those in the private sector was not so high.

ii) Public disclosure of evaluation results

[Extent of disclosure]

Analysis revealed specific trends by organization (ministries and agencies). The Science and Technology Agency ranged from 20 to 40 pages per institute averaging 21.5 pages for the volume of the evaluation report. In the case of the Ministry of Education, the National Astronomical Observatory came up with 26.5 pages and the Okazaki National Research Institute with 32.9 pages, averaging 20.2 pages. In the case of the Ministry of Education, however, only three organizations, the National Astronomical Observatory, the Okazaki National Research Institute, and the National Aerospace Laboratory of Japan disclosed the results of evaluation out of the 14 joint college research institutes administered by the Ministry. This also holds true for the Ministry of Health and Welfare, where it was only the National Children's Hospital (National Children's Medical Research Center) that disclosed the results of evaluations of about 15 pages, with 12 other institutes neglecting public disclosure. On the other hand, a fairly large number of institutes (19 out of 29 institutes) under the Ministry of Agriculture, Forestry and Fisheries disclosed the results of evaluations. However, the extent of disclosure was small, averaging about five pages per institute. Under the Ministry of International Trade and Industry, 8 out of the 15 institutes disclosed evaluation results in considerable detail, with disclosure per institute averaging 49 pages, far ahead of other ministries and agencies. Further, the National Institute of Telecommunications of the Ministry of Posts and Telecommunications came up with moderate disclosure of about 7.5 pages, the Ship Research Institute of the Ministry of Transport about 3 pages, and the National Institute of Industrial Safety with a brief summary in a single page.

Thus, the Science and Technology Agency and the Ministry of International Trade and Industry have advanced to a positive disclosure of the results of evaluations. Most institutes under the Ministry of Education and the Ministry of Health and Welfare failed to disclose the results of evaluations, except a few institutes which were disclosed in fairly good detail. On the other hand, many institutes under the Ministry of Agriculture, Forestry and Fisheries took a positive stance towards disclosure, but the extent of disclosure per institute was small.

[Forms of evaluation rating]

A majority of institutes presented evaluation reports as descriptive ratings, with the exception of only three institutes presenting them in gradation ratings, including the National Research Institute for Earth Science and the Disaster Prevention of the Science and Technology Agency, the National Children's Hospital (National Children's Medical Research Center) of the Ministry of Health and Welfare and the National Institute of Industrial Safety of the Ministry of Labor.

[Contents of evaluation results]

All institute evaluation reports were presented as descriptive ratings. An analysis was not performed due to the complexity and obscurity of descriptions in the reports.

iii) Utilization of evaluation results

In all, 18 research institutes disclosed policies and measures to put into effect the results of evaluation. Of these, there were positive disclosures by 13 institutes under the Ministry of Agriculture, Forestry and Fisheries, but as to the state of utilization, few cases of description was seen as it was shortly after the evaluation.

2.2 R&D evaluation by internal subdivisions

Research and development schemes carried out by the internal subdivisions of ministries and agencies are classified into those implemented under the research and development system and those implemented on a project basis. Under the research and development system, multiple projects are implemented individually and evaluation is performed by project. On the other hand, research and development carried out on a project basis covers large-scale R&D projects, and evaluation is applied to the project as a whole. Accordingly, as to the evaluation of research and development conducted by internal subdivisions, discrete analysis is made of the research and development system and the research and development project.

The scope of analysis includes nine items relating to expenditure on promotion and adjustment of science and technology, regional research and development promotion schemes and computing science and technology joint research promotion schemes of the Science and Technology Agency; expenditure on the promotion of basic research for creation of future environment and expenditure on research of anti-pollution technology, etc. by the national research institutes of the Environment Agency; project research⁷⁾ of the Ministry of Agriculture, Forestry and Fisheries; and industrial science and technology research and development schemes, comprehensive energy and environment technology development schemes (New Sunshine Project), and a key regional technology research and development scheme of the Ministry of International Trade and Industry.

Also, research and development implemented on a project basis relates to an analysis of nine items covering the phase II superconductive material research multicore project, the earth simulator plan, the deep-sea earth drilling plan, research and development concerning advanced supersonic aircraft technology of the Science and Technology Agency; research and development for the stratosphere platform jointly steered by the Science and Technology Agency and the Ministry of Posts and Telecommunications; the gene bank project undertaken by the Ministry of Agriculture, Forestry and Fisheries; development of energy-saving metal manufacturing processes, research and development concerning soil de-contamination technology by the Ministry of International Trade and Industry; and research and development for practical technology on MAGLEV vehicles propelled by linear motor as steered by the Ministry of Transport.

(1) Formulation and disclosure of evaluation procedures

Evaluation procedures were laid down and disclosed under the research and development system with respect to “expenditure on promotion and adjustment of science and technology” and “promotion of joint research on computing science and technology” of the Science and Technology Agency and the project research of the Ministry of Agriculture, Forestry and Fisheries. Also, evaluation procedures were laid down and disclosed under the research and development project with respect to the deep-sea earth drilling plan of the Science and Technology Agency, and the gene bank project of the Ministry of Agriculture, Forestry and Fisheries.

Thus, with respect to R&D evaluation conducted by internal subdivisions, little progress has been made in the formulation and disclosure of evaluation procedures. In the case of most of the other schemes and projects, evaluation reports carry basic entries on performance of evaluation, but it is nevertheless difficult to find details regarding the method of implementing the evaluations.

(2) Progress in the introduction of external evaluation

[Acceptance of outside evaluators]

7) Practically speaking, there is no research and development system called “project research,” but in the case of the Ministry of Agriculture, Forestry and Fisheries, the Conference on Agriculture, Forestry and Fisheries Technology evaluates large-scale R&D schemes implemented under their research and development system collectively as project research. Accordingly, this paper represents them collectively as project research.

Under the research and development system, most projects and schemes were evaluated by outside evaluators (Figure 8), with the exception of project research by the Ministry of Agriculture, Forestry and Fisheries where internal staff joined the evaluation. Internal staff also joined the evaluations conducted by the Environment Agency relating to expenditure on the promotion of basic research for the creation of future environment and expenditure on research concerning anti-pollution technology, etc. by national research institutes. It should be noted that the bar graph was converted into graduation since the Environment Agency had not disclosed the number of internal evaluators. Here, "internal evaluators" point to personnel from the internal subdivisions of ministries/agencies and from national research institutes.

More evaluators were assigned to the evaluation of expenditure on the promotion and adjustment of science and technology incurred by the Science and Technology Agency and project research carried out by the Ministry of Agriculture, Forestry and Fisheries. How many evaluators are involved per project? As many as seven evaluators were employed for the evaluation of each scheme conducted by the Ministry of Agriculture, Forestry and Fisheries and the Ministry of International Trade and Industry, compared with as few as four evaluators employed for the evaluation of expenditure on the promotion and adjustment of science and technology incurred by the Science and Technology Agency. This indicates that the Ministry of Agriculture, Forestry and Fisheries and the Ministry of International Trade and Industry are more aggressive in providing evaluators per project.

On the other hand, most of the evaluation of "research and development of practical technology for MAGLEV vehicles propelled by linear motors" by the Ministry of Transport was carried out by outside evaluators although a few internal staff joined the evaluation.

[Composition of outside evaluators]

Few foreign evaluators, if any, were employed for the evaluation of R&D projects under the research and development system, with the single exception of the Science and Technology Agency, where a Thai university professor joined the evaluation of a project relating to expenditure on the promotion and adjustment of science and technology.

(3) Disclosure and contents of evaluation

[Extent of disclosure]

Under the research and development system, a large volume of evaluation results were disclosed relating to "expenditure on the promotion and adjustment of science and technology" of the Science and Technology Agency; "the project research" of the Ministry of Agriculture, Forestry and Fisheries; "research and development of industrial science and technology" and "the New Sunshine Project" of the Ministry of International Trade and

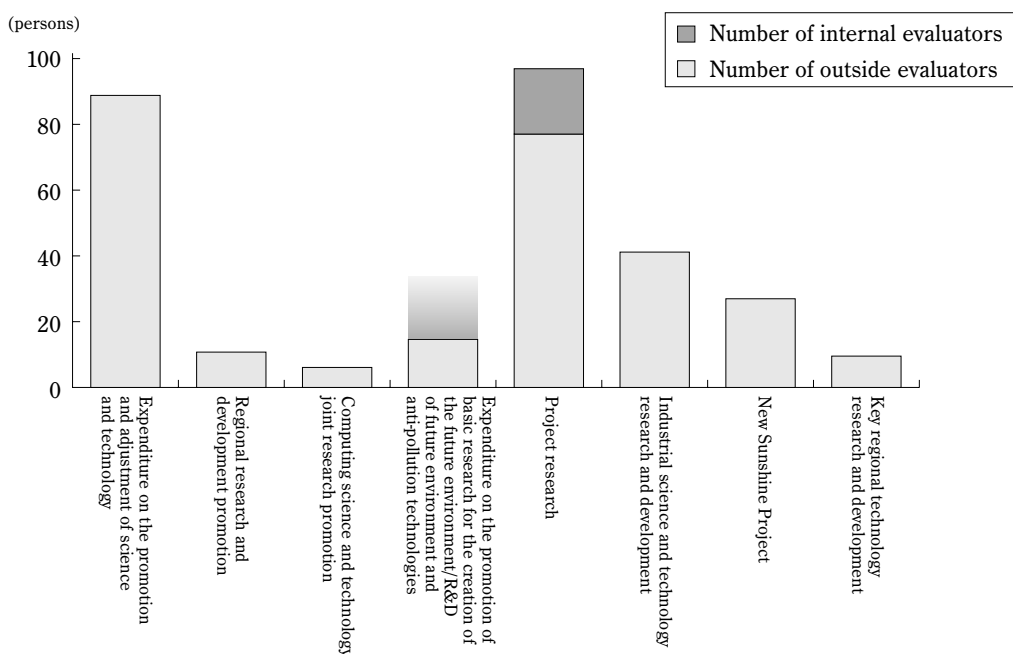


Figure 8 Number of Internal and Outside Evaluators (Research and Development System)

Industry. In terms of the extent of disclosure, the Ministry of International Trade and Industry topped other ministries and agencies by 2 to 3 pages per project.

As to the Research and Development Project, the Science and Technology Agency describes even the gist of minutes in the results of evaluation relating to research and development of advanced supersonic aircraft technology (50 pages in all), much ahead of others.

[Forms of evaluation ratings]

Most ministries and agencies state the results of evaluations as descriptive ratings, while the Science and Technology Agency exceptionally presented them in gradation ratings with respect to the evaluation of "expenditure on the promotion and adjustment of science and technology" and "the promotion of joint research on computing science and technology." Notably, the gradation ratings applied to evaluations of "expenditure on the promotion and adjustment of science and technology" is quite clear and comprehensible. Results are rated in seven gradations of "a, a-, b+, b, b-, c+, c" and are supplemented by detailed descriptive ratings. Preferably, other research institutes should use such concise and comprehensible rating method as far as possible.

Then, in the case of research and development projects, evaluation reports, mostly presented as descriptive ratings, were less comprehensible. Gradation ratings, if any, were applied solely to item-wise evaluations, but not to comprehensive project evaluations.

[Rates of pre-, interim- and post-evaluation disclosures]

It turned out that interim and post-evaluation disclosure rates were high under the Research and Development System, while the rate of interim-evaluation disclosure was high under Research and Development Projects. It is assumed that pre-evaluation has been conducted by the Science and Technology Agency on "expenditure for the promotion and adjustment of science and technology", but the results of the evaluations have not been undisclosed, for reasons that the author cannot identify.

[Results of evaluation]

Most evaluation reports are presented as descriptive rating as mentioned, which is often responsible for the complexity and incomprehensibility of the contents involved. As a result, it is extremely difficult to rate them as "high," "fair on the whole or ordinary" or "low" gradations. Exceptionally, such gradation ratings could be applied to the Science and Technology Agency's clear-cut descriptive evaluation rating on expenditure on promotion and adjustment of science and technology. As a result of rating them by gradation method, 85% were rated "high," 5% "ordinary," and 10% "low." Thus, the majority of project evaluations were rated "high." Among them, few projects were rated "low," but this is a rare case.

(4) Utilization of evaluation results

The Ministry of Agriculture, Forestry and Fisheries earlier presented a policy on the research plan and system, while other ministries and agencies did not go further than the disclosure of evaluation results.

3. Integration of Indicators for Governmental R&D Evaluation Activities

By looking at the integration of indicators of results of analysis from four discrete viewpoints, this paper analyzes the overall trend of research and development by ministries and agencies, where multivariate analysis (main component analysis) method was adopted as a technique suitable for indicator integration. For the variables of the multivariate analysis, conventional indicators are used.

- Evaluation procedures

Formulation of evaluation procedures (1 if formulated, 0 if otherwise) / Disclosure of evaluation procedures (1 if disclosed, 0 if otherwise)

- External evaluation

Number of evaluators: Number of evaluators per project (not applicable to institute evaluation)/number of outside evaluators/number of foreign evaluators/number of evaluators from the private sector/number of evaluators from the academic sector/number of evaluators from the administrative sector/number of evaluators from the public service corporation/number of evaluators from competent ministries or agencies/number of

internal evaluators

- Disclosure of evaluation results

Extent of disclosure/extent of disclosure per project (not applicable to institute evaluation)/number of projects evaluated (not applicable to institute evaluation)/number of pre-evaluated projects (not applicable to institute evaluation)/number of interim-evaluated projects (not applicable to institute evaluation)/number of post-evaluated projects (not applicable to institute evaluation)/comprehensive evaluation (2 for parallel use of descriptive and gradation ratings, 1.5 for exclusive use of gradation rating, 1 for exclusive use of descriptive rating and 0 for none use) /item-wise evaluation (2 for parallel use of descriptive and gradation ratings, 1.5 for exclusive use of gradation rating, 1 for exclusive use of descriptive rating and 0 for non-use)

- Utilization of evaluation results

Countermeasures for evaluation (1 if established, 0 if otherwise)

3.1 Project Evaluation for National Research Institutes

(1) Interpretation of main components

Analysis was made up to the second main components, whose cumulative contribution rate was 48.7%. Each component is mutually independent.

In terms of vectors specific to the first component, “number of evaluators from the private sector,” “number of evaluators from the administrative sector,” “number of evaluators from the academic sector,” “extent of evaluation disclosure,” etc. are distributed on the right side of Figure 9, while “number of internal evaluators,” “number of pre-evaluations,” “number of projects evaluated” and “number of post-evaluations” are distributed on the left side of the figure. Accordingly, the presence of a high value for the first main components indicates the active introduction of outside evaluators and the disclosure of evaluation results. On the other hand, the presence of a low value indicates active evaluation of numerous projects irrespective of the introduction of outside evaluators. This can be defined as a “relation of emphasis on quality versus emphasis on quantity.”

In terms of vectors specific to the second main components, there are many factors that indicate the overall quantity, i.e. “extent of evaluation disclosure,” “number of evaluators,” “number of foreign evaluators,” “comprehensive evaluation,” “interim-evaluation” and “number of projects evaluated,” specifically representing the level of evaluator input (irrespective of internal or external), projects evaluation and results disclosure. Therefore, the second main components can be defined as the level of labor input for evaluation.

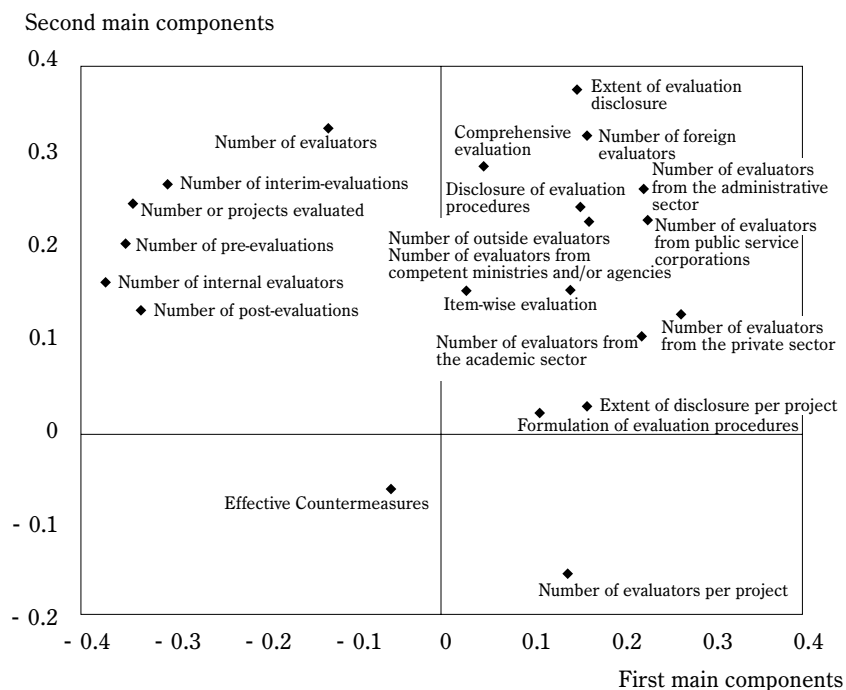


Figure 9 Distribution of Specific Vectors

(2) Main component scores

Figure 10 represents the distribution of the first and the second component scores of respective national research institutes. Institute with high scores of the first and the second main components and that placing an emphasis on the factor of quality of evaluation and high labor input was the National Institute of Radiological Sciences of the Science and Technology Agency. On the other hand, the institutes laying emphasis on the factor of quantity were the National Institute of Livestock and Agrobiological Sciences and the Tohoku National Fisheries Research Institute of the Ministry of Agriculture, Forestry and Fisheries. Also, the National Research Institute for Metals of the Science and Technology Agency features a high level of labor input. Most of other institutes show a dominant distribution of the first and the second main component scores in a range from -2 through 2, showing inter-institute similarity in respect to quality, quantity, and level of labor input in evaluation.

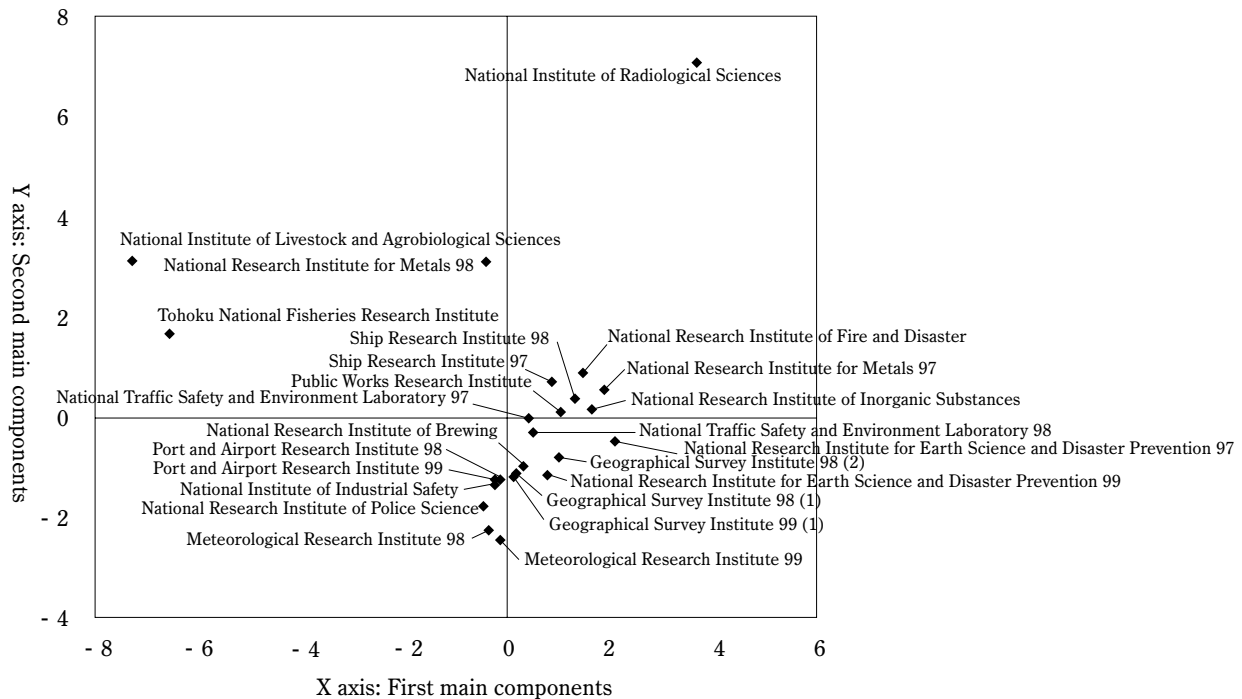


Figure 10 Distribution of Main Component Scores

3.2 Institute Evaluation of National Research Institutes

(1) Interpretation of main components

Analysis was made up to the second main components concerning institute evaluation, whose cumulative contribution rate was 47.7%.

In terms of vectors specific to the first main components, factors of “number of outside evaluators,” “number of evaluators,” “number of evaluators from the academic sector” and “number of evaluators from the administrative sector” showed high values, while the factors of “number of internal evaluators” and “number of foreign evaluators” showed low values (Figure 11). From these findings, the author concludes that the first main components represent the level of input of Japanese outside evaluators.

In terms of vectors specific to the second main component, the factors of “number of foreign evaluators,” “disclosure of evaluation procedures,” “extent of disclosure” and “formulation of evaluation procedures” etc. show high values, while the factors of “disclosure of countermeasures,” “number of evaluators from competent ministries and agencies,” “number of evaluators from the administrative sector,” etc. show low values. It appears that the second main components represent a relation between the “type in which emphasis is placed on processes leading to the presentation of evaluation results,” such as active formulation of evaluation procedures and the disclosure of evaluation results, and the “type in which emphasis is placed on how to take effective countermeasures following evaluation.”

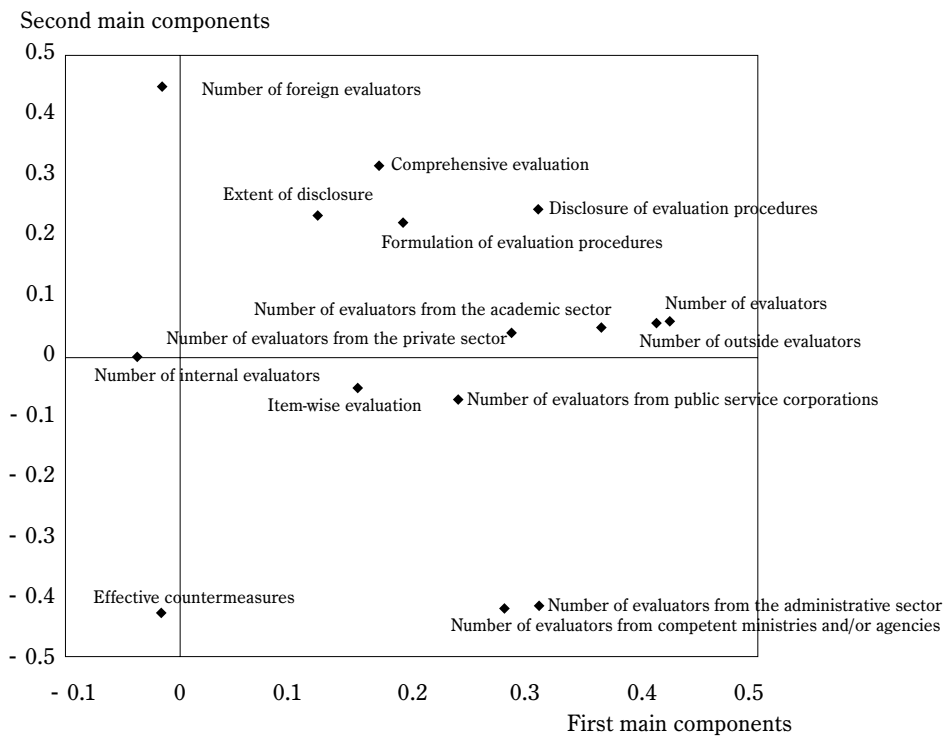


Figure 11 Distribution of Specific Vectors

(2) Main component scores

Figure 12 represents the distribution of the scores of the first and the second main components applicable to the respective national research institutes. Institutes showing high values in the first main components (representing the input level of Japanese outside evaluators) are National Aerospace of Japan under the Science and Technology Agency, Okazaki National Research Institute and National Children's Hospital (National Children's Medical Research Center) of the Ministry of Education. Other institutes show dominant distribution in a range

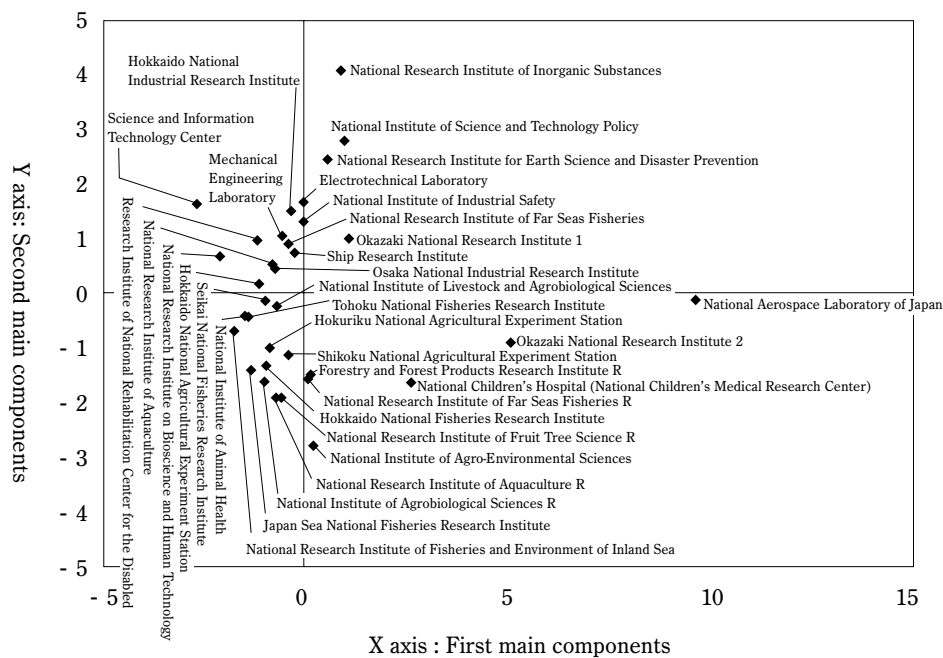


Figure 12 Distribution of Main Component Scores

from -2 through 1 showing an absence of any major difference in the level of output of Japanese outside evaluators.

Research institutes under the Science and Technology Agency show a dense distribution in the upper region of the figure, which represents the “type in which emphasis is placed on processes leading to presentation of evaluation results.” On the other hand, counterparts of the Ministry of Agriculture, Forestry and Fisheries are distributed more in the lower region of the figure which represents the “type in which emphasis is placed on how to take effective countermeasures after evaluation.” The intermediate area in a range between 0 and 2 shows a dense distribution of research institutes of the Ministry of International Trade and Industry. Many research institutes under the Ministry of International Trade and Industry featured aggressive introductions of foreign evaluators and substantial disclosure of evaluation results, including the provision of effective countermeasures. They feature an “intermediate type” between the former and the latter types. Thus, respective ministries and agencies can be classified into the “type in which emphasis is placed on processes leading to presentation of evaluation results” and the “type in which emphasis is placed on how to take effective countermeasures after evaluation,” and the “intermediate type.”

3.3 Internal Subdivisions Project Evaluation

(1) Interpretation of main components

Analysis was made up to the second main components, whose cumulative contribution rate was 61.4%.

In terms of vectors specific to the first main components, the factors of “number of outside evaluators,” “number of evaluators,” “number of evaluators from the academic sector,” “Number of evaluators from the private sector,” “number of projects evaluated,” etc. show high values, while the factors that showed low values were “extent of disclosure per project,” “number of evaluators per project,” etc. (Figure 13). In other words, most of the factors showing higher values represented total number of evaluators and projects evaluated, while those showing lower values represent the number of evaluators and extent of disclosure per project. Accordingly, the first main components, the author interprets, can be defined as a correlation of “total value” and “per project.”

In terms of vectors specific to the second main components, the factors of “number of internal evaluators,” “number of pre-evaluated projects,” etc. show higher values, and the factors of “number of post-evaluated projects,” “number of projects evaluated,” “number of interim-evaluation projects,” etc. show lower values. The author defines the second main components as representing correlation of a pre-evaluation type versus interim- and post-evaluation types since the “number of pre-evaluations” relates oppositely to the “number of post-evaluations” and the “number of interim-evaluations.”

(2) Main component scores

Figure 14 shows the distribution of the scores for the first and the second main components applicable to respective R&D systems and projects. Among the first main components, “expenditure on the promotion and adjustment of science and technology” of the Science and Technology Agency and “project research” of the Ministry of Agriculture, Forestry and Fisheries show high values. These are covered by the R&D system and involve numerous projects to be evaluated. Items covering "expenditure on research of anti-pollution technology applicable to the national research institutes " and "expenditure on the promotion of basic research for creation of the future environment" of the Environment Agency, "research and development of the industrial science and technology " and "the New Sunshine Project" of the Ministry of International Trade and Industry assume positive value in the first main component scores, specifically assuming high total values in “number of evaluators” and “number of projects.” On the other hand, most research and development projects consisting of a single theme show negative values.

In terms of the second main component, the upper area of Figure 13 shows the dense distribution of projects under the auspices of the Ministry of Agriculture, Forestry and Fisheries, indicating the increasing trend of disclosing pre-evaluations. On the other hand, the lower area of the figure shows the dense distribution of items covering "expenditure on the promotion and adjustment of science and technology" applicable to the Science and Technology Agency, and "expenditure on research of anti-pollution technology" and "expenditure on the promotion of basic research for creation of future environment" applicable to the Environment Agency and "the research and

development system" of the Ministry of International Trade and Industry. Specifically, the lower area shows a more dense distribution of research institutes conducting interim- and post-evaluations.

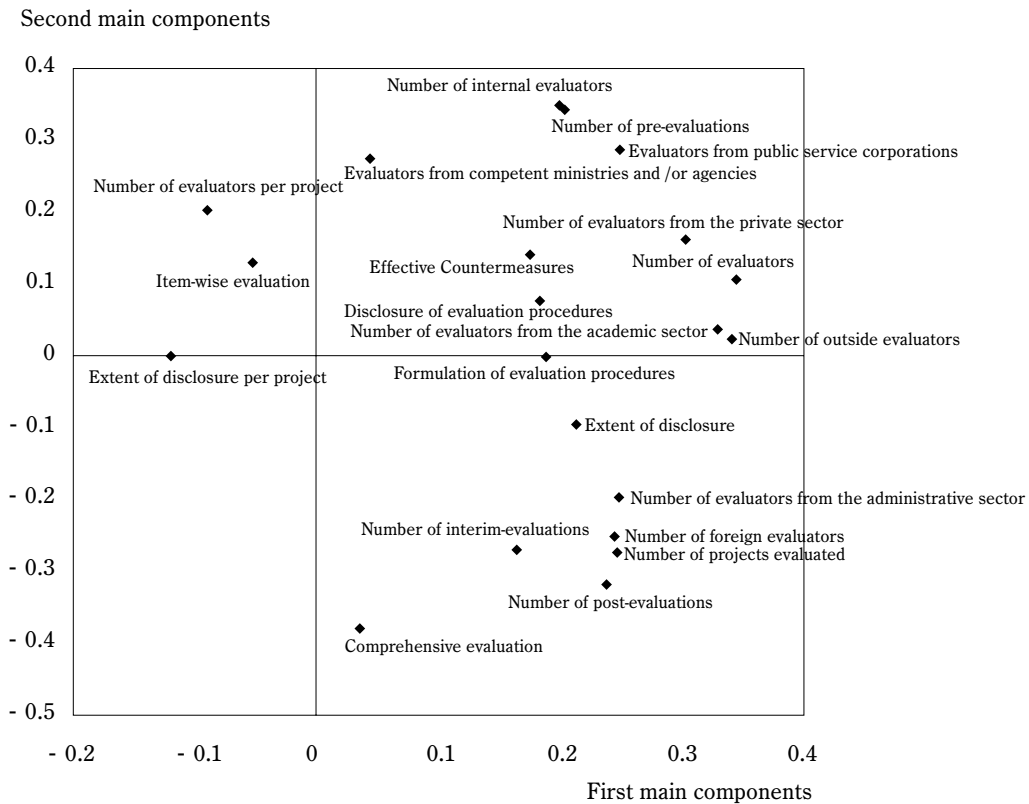


Figure 13 Distribution of Specific Vectors

4. Conclusion-Characteristics and Prospects of R&D Evaluation of Ministries and Agencies-

4.1 Characteristics of Evaluation

The paper presents the characteristics of research and development carried out by relevant ministries and agencies based on the results of analysis.

National Police Agency

Research and development applicable to the National Police Agency has been carried out by National Research Institute of Police Science. Although there is slight lack of transparency concerning evaluation procedures and outside evaluators, the results of evaluation are supported by concise but comprehensible descriptions and the contents have been disclosed on the Internet Web site. No mention was made of countermeasures concerning the evaluations, nor the current status of utilization of recommendations.

Science and Technology Agency

As a leading state organization playing a central role in the introduction of the R&D evaluation system, the Agency's research institutes and internal subdivisions are taking a positive stance towards the formulation and disclosure of evaluation guidelines, the introduction of outside evaluators and Internet disclosure of evaluation results. However, it appears that plans to utilize recommendations and results of utilization thereof were sparingly disclosed.

Environment Agency

The Environment Agency runs two national research institutes, but no information on evaluation was disclosed on their Internet Web site. Internal subdivisions disclosed them on two systems, but the descriptions were brief, which impeded closer access to the contents of R&D evaluation as a whole.

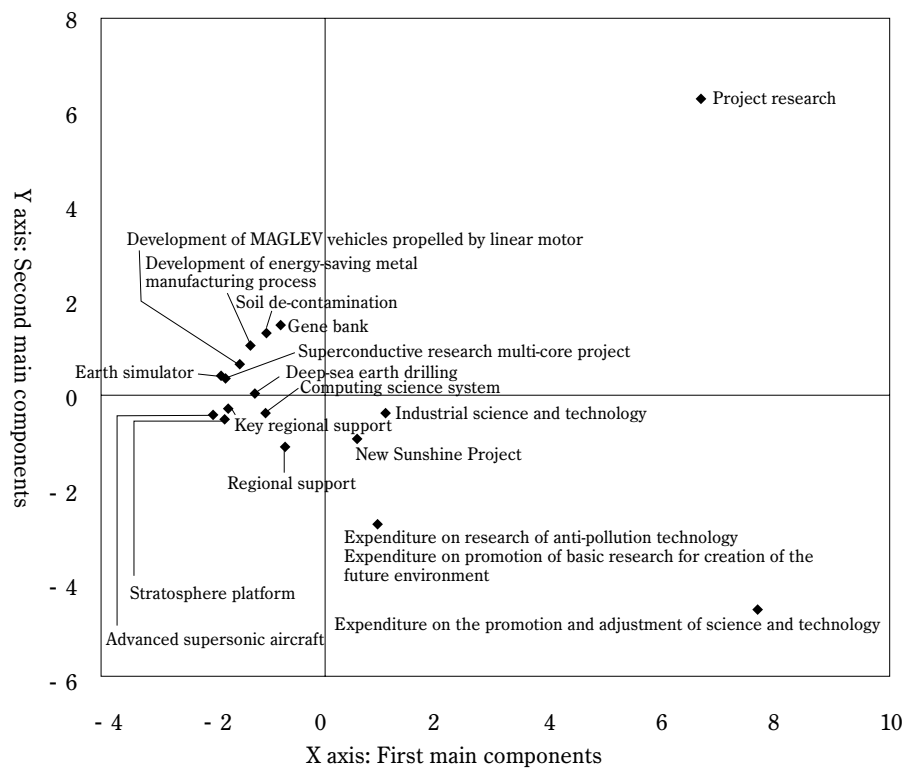


Figure 14 Distribution of main component scores

Ministry of Finance

Research and development has been carried out by the National Research Institute of Brewing under the National Tax Administration Agency. Evaluation procedures were laid down, but not disclosed. Outside evaluators have been introduced, with their names and position titles disclosed. The contents of evaluation are quite concisely described. As to the utilization of evaluation results, amendments have been made to the R&D plans, but relevant details remain undisclosed.

Ministry of Education

The Ministry of Education's research institutes (referring to joint college research institutes) are not taking a positive stance toward disclosure of evaluation information on the Internet. Therefore, it is difficult to discover how evaluation has been practiced and how the outcome of research and development has been utilized. This also holds true for the evaluation of research and development by internal subdivisions. Since the Ministry of Education enjoys a more liberal R&D budget allocation than its counterparts, it is desirable that the Ministry takes active steps concerning disclosure. Notably, since the University Evaluation and Degree Grant System was instituted in FY 2000, followed by the prospect for full-scale evaluation of the joint college research institute, high hopes are pinned on future achievements.

Ministry of Health and Welfare

The Ministry administers 13 national research institutes. Of them, it was only the National Children's Hospital (National Children's Medical Research Center) that disclosed evaluation results. However, the Center did not disclose evaluation procedures, and internal evaluators also joined the Evaluation Committee as committee members. Evaluation results, presented as gradation ratings by division, are clear and transparent.

The Center also forged countermeasures for the results of evaluations. In the future, 12 other research institutes are expected to undertake positive disclosure. However, in view of the fact that each institute is operating on a limited budget, it is recommendable for these institutes to conduct evaluation commensurate to the scale of their budgets.

Further, less Internet disclosure was made of the results of evaluation by internal subdivisions.

Ministry of Agriculture, Forestry and Fisheries

At the Ministry of Agriculture, Forestry and Fisheries, the Technological Conference on Agriculture, Forestry

and Fisheries plays a central role in evaluation, where significant Internet disclosure has been made of evaluation information. This holds true with evaluation of research institutes, research and development projects systems and projects as a whole.

Nevertheless, a trend was observed where internal evaluators and those close to internal organizations joined the evaluations. The Ministry disclosed countermeasures concerning evaluation results in more detail than its counterparts. The Ministry lays emphasis on effective use of recommendations based on evaluations as well as on the evaluations themselves.

Ministry of International Trade and Industry

Under the Ministry of International Trade and Industry, the Agency of Industrial Science and Technology has been playing a central role in evaluation since it established the Technology Evaluation Division in FY 1997. Each institute has been involved in institute evaluation, disclosing relevant results in an active manner. Institutes under the Ministry disclosed a fairly large volume of detailed evaluation results. Further, the Ministry appears to pay careful attention to the deployment of outside evaluators in reasonably neutral positions and establish reasonable countermeasures for evaluations.

However it was often impossible to acquire clear information concerning the introduction of evaluation procedures, although it appears that each institute is conducting evaluations based on the ministry-level technology evaluation guidelines. As to evaluation results, most of them were presented as descriptive ratings. In this connection, the author would hope that they would be presented in the forms of gradation ratings in the future.

Positive disclosure was made of results of project evaluations on the research and development system and other projects such as "research and development of the industrial science and technology " and "the New Sunshine Project", though these were not accompanied by descriptions of countermeasures and recommendations based on the evaluations.

Ministry of Transport

Project evaluation has been implemented by national research institutes under the Ministry of Transport, followed by positive disclosure of relevant results. Evaluation guidelines and procedures have been laid down by most institutes, including Internet disclosure by some of them. The introduction of outside evaluators has made some progress. However, a large part of evaluation ratings were represented by the abstract phrase of "fair on the whole." In the future, it should be amended to state the ratings in a more concrete, comprehensible way. Also, no reference was made to countermeasures for recommendations based on the evaluations or to the status of the utilization thereof.

Sparing disclosure was made of the results of R&D evaluations by internal subdivisions.

Ministry of Posts and Telecommunications

Research and development has been carried out by the National Research Institute of Telecommunications under the Ministry of Posts and Telecommunications. These institutes implemented evaluations prior to the introduction of the National Guidelines on Evaluation, including active introduction of outside evaluators and Internet disclosure of evaluation results.

For evaluation of R&D projects by internal subdivisions, outside evaluators were employed and this was accompanied by disclosure of evaluation results.

Ministry of Labor

Research and development has been carried out by the National Institute of Industrial Safety and National Research Institute of Industrial Medicine under the Ministry of Labor. Internet disclosure of evaluation results was made only by the former organization, but the contents disclosed were no more than brief summaries. Moreover, no disclosure was made of evaluators' names and position titles.

Ministry of Construction

The results of project evaluations have been disclosed by the Public Research Institute and the Geographical Survey Institute among the three research institutes under Ministry of Construction. No disclosure has been made of evaluation procedures on the Internet. No information was obtained even on their formulation. As to the introduction of outside evaluators, Public Works Research Institute assigned outside evaluators to the evaluation of three projects they considered important out of eight projects. The Public Works Research Institute then

established plans for effective countermeasures following the evaluation.

As to research and development evaluation by internal subdivisions, no information was disclosed on the Internet.

Ministry of Home Affairs

Research and development has been carried out by the National Research Institute of Fire and Disaster of the Fire Defense Agency under the Ministry of Home Affairs. Evaluation procedures have been formulated and disclosed. The Institute introduced outside evaluators and disclosed them, including their names and position titles. The Institute also provided detailed description of evaluation results, though unaccompanied by any description of countermeasures for the evaluation recommendations. Though operating on a moderate scale, the Institute has been doing well in performance of evaluation and the disclosure of relevant results.

4.2 Recommendations for Future R & D Project Evaluation

Based on the foregoing analysis, the paper presents the following suggestions concerning the evaluation of R&D projects in the future.

(1) Positive disclosure of evaluation procedures

In the cases of more than half of all national research institutes, it was impossible to determine if they were operating on their own evaluation procedures or on the ministry-level evaluation guidelines. Fewer institutes disclosed their evaluation procedures on the Internet. Unless explicit statement is made of the criteria or arrangements that each evaluation is based on, this will surely arouse doubts or questions on given evaluation results, for example relating to the question of how outside evaluators were selected, how far the scope of evaluation extends, etc. Unless such factors are made clear, it is impossible to accurately judge any given evaluation results for reliability, equity, etc. even when they are disclosed.

As the National Guidelines on Evaluation states "It is imperative to introduce a transparent and objective evaluation methodology to enable the public to understand the contents of evaluations", the author believes it necessary to take positive steps towards the extensive disclosure of evaluation guidelines as well as results of evaluation.

(2) Selection of outside evaluators of neutral positions

Upon analysis of outside evaluators by origin (sector), it turned out that some come from the ministries/agencies administering the research institutes that were employing them as evaluators, or from public service corporations affiliated to the research institutes employing them as evaluators. Thus, the possibility cannot be ruled out that personnel working as outside evaluators of given research institutes are retired officers of these research institutes or of ministries or agencies governing these institutes. Such selection practice for outside evaluators can obscure or complicate their neutrality towards the interests of the research institutes evaluated by them. Accordingly, it is imperative to select neutral outside evaluators under the objective evaluation procedures and disclose these procedures to the public. In reality, however, the selection of completely neutral outside evaluators is a complex and time consuming issue and therefore, it is preferable to balance the selection of qualified evaluators with the efficiency of selection.

(3) Comprehensible description of evaluation results

In the case of evaluation of research and development, a description of evaluation results in a technical style can baffle ordinary readers, while if they are presented with a gradation rating, as of an "ABC-type rating," people without expertise can approximately check the research works for propriety. In this connection, the author pins hopes on the increasing introduction of clear, comprehensible rating formulae, like those applied to the evaluation of "expenditure on the promotion and adjustment of science and technology" and to the evaluation of projects by the medical and related divisions of the National Institute of Radiological Sciences of the Science and Technology Agency.

(4) Reports on formulation and utilization of countermeasures concerning evaluation

Upon the perusal of evaluation reports, it is natural that readers should be interested in how the results of evaluation can be or have been utilized. Unexpectedly, few institutes presented ways for effective utilization, despite disclosing their evaluation results. The few exceptions included the Mechanical Engineering Laboratory of the Ministry of International Trade and Industry, which introduced Internet disclosure of their evaluation results, setting a link to the "Action Plan of Mechanical Engineering Laboratory" on the screen. When clicked, the screen displays the Laboratory's action plans showing countermeasures for the evaluation presented. This gives searchers easy access to ways concerning the effective use of evaluation results.

Such practices in the disclosure of countermeasures are likely to be followed by other ministries and agencies. For example, the Ministry of Agriculture, Forestry and Fisheries has been taking an active stance towards the disclosure of their policies and countermeasures though it did not go as far as setting up links to them.

Disclosure of status on the utilization of evaluations is rare. Of course, people should not be blamed for the failure to report the status of utilization promptly following the results of an evaluation. Nevertheless, more efforts should be made from now on to also disclose the state of utilization.

(5) Statement of evaluation costs

Evaluation costs manpower, money, and time. Should it cost too much, it could impede the fair progress of research and development itself. Also, it is often pointed out that extra assignments concerned with evaluations can deprive researchers of the time to devote on their major research duties, with much of their time taken up by extra missions as evaluators. On the other hand, it is true that spending time and labor on the task of evaluation, for example, the preparation of reference materials for evaluation, is conducive to upgrading their research. In any case, the cost issue weighs heavily on the task of R&D evaluation.

In the future, ministries and agencies will be required to determine the actual costs on evaluation and initiate discussions on how to determine the optimum level of costs.

(6) From project-level to program-level evaluation

As discussed in this paper, national research and development evaluation generally belongs to the "project" area, in terms of the three level classification of "policy," "program," and "project." In the future it will be necessary to discuss how to implement "program" level evaluation. At present, if the research institute is defined as a type of "program" aimed at implementing each R&D project, "institute evaluation," which evaluates the policy and management of research institutes, can also be defined as "program" level evaluation.

With respect to the research and development system, however, an evaluation is made of each project carried out under the system, but not the system itself. The author thinks that attention should also be directed at system evaluation as well.

As mentioned earlier, this paper makes it clear that evaluation practices vary with the institute and that some of them undertook sufficient evaluation. The paper characterized the specific evaluations of each ministry and agency through a parallel, comparative analysis of these organizations in terms of the evaluation of research and development and recommends how evaluation should be carried out in the future. It should be noted that national research and development has just made a new start and since there is no decisive method for implementing the evaluation of R&D projects, each government organization will be required to go through a variety of "try and error" processes on its own. At this crucial stage, boldly facing the challenges to things new is essential. This can lead to the effective and efficient evaluation of R&D projects and so to higher level of research and development. The author concludes this paper by hoping that it is instrumental in furthering this process.

The author would like to add that this paper focused its analysis on the evaluation of research and development projects by ministries and agencies applicable to the period of June 16 through November 15, 1999 (five months), and that the contents may in part depart from the present situation following the lapse of time since then.

Bibliography

- “Multivariable Analysis Method” authored by Chuichi Okuno, Hitoshi Kume, Toshiro Haga and Masashi Yoshizawa (Japan League of Science and Technology, 1971)
- “Current Status of R&D Evaluation” edited by Science and Technology Agency (Printing Bureau, Ministry of Finance, 1999)
- “Investigation into Ways for R&D Evaluation” edited by the R&D Evaluation Techniques Investigation Committee (Asahi Research Center, 1984)
- “Message Analysis Techniques” authored by Claus Krippendorf and translated by Toshiharu Mikami, Nobuo Shiino, Yoshiaki Hashimoto (Keiso Shobo, 1989)
- “R&D Projects Evaluation Techniques Investigation Report” edited by Institute of Policy Science (Institute of Policy Science, 1999)
- “Policy Evaluation Techniques” edited by Association of New Administrative System Research (Articles carried on the Tokyo District morning editions of April 14 through May 12, 1999 of Nihon Keizai Shimbun)
- “Analysis of Current Developments in R&D Evaluation by Ministries and Agencies” authored by Fujio Niwa and Keita Yamazaki (Gist of lectures delivered at the 14th Annual Meeting of the Council of Research and Technology Planning, 1999)
- “Strategic R&D Evaluation and Decision Making” edited by the Planning and Editing Committee on “R&D Evaluation and Decision Making” of Japan Management Association (Japan Management Association, 1982)
- “Theory of Policy Evaluation and its Development” authored by Kiyoshi Yamaya (Koyo Shobo, 1997)
- “External and Trade Association and Advisory Bodies of Ministries and Agencies” authored by Kyoko Kasa in “Lectures on Public Administration” (Volume 4) edited by Masaru Nishio and Michio Muramatsu (Yuhikaku, 1994)
- AAAS Science and Technology Policy Yearbook 1999
- Homepages of ministries and agencies of Japanese Government