

tentative version

National Guidelines for Evaluating Government Funded R&D

March 29, 2005

The Prime Minister of Japan

(Table of contents)

Introduction	3
Chapter 1: Basic Concept and Framework Of R&D Evaluation	7
1. Significance of Evaluation	7
2. Scope of Application of these Guidelines	8
3. Responsibilities of Evaluating Organizations, Evaluators and Researchers	8
(1) Responsibilities of R&D Implementation/Promotion Organizations	8
(2) Responsibilities of Evaluators	9
(3) Responsibilities of Researchers	9
4. Direction of Evaluation System Reform	10
5. Follow-Up on the Implementation of These Guidelines	11
Chapter 2: Common Principles in Conducting Evaluation	12
1. Setting of Evaluation Subjects	12
2. Setting of Evaluation Objectives	12
3. Selection of Evaluators	13
4. Setting of Evaluation Timing	15
5. Setting of Evaluation Methods	17
(1) Evaluation Technique	17
(2) Evaluation Perspectives	19
(3) Evaluation Items and Criteria	20
(4) Setting of Flexible Evaluation Methods	21
(5) Averting Excessive Burdens in Evaluation	23
6. Utilization and Handling of Evaluation Results	24
(1) Utilization of Evaluation Results	24
(2) Disclosing Evaluation Contents to Evaluated Parties	26

(3) Publishing R&D Evaluations	27
7. Administering an Effective and Efficient Evaluation System...	28
(1) Administering Evaluations in a Hierarchical Structure	29
(2) Administering Chronological Evaluations	30
(3) Review of Evaluation System	31
8. Improvement of Evaluation Implementation System	32
(1) Developing and Securing Human Resources for Evaluation and Raising the Level of Evaluations	32
(2) Database Development and Introduction of Electronic Systems for Efficient Evaluation	34
Chapter 3: Facts to Pay Attention to, with Different Evaluation	
Subjects	36
1. Evaluation of R&D Measures	36
2. Evaluation of R&D Themes	37
(1) Competitive Fund Themes	37
(2) Priority Fund Themes	39
(3) Base Fund Themes	39
3. Evaluation of R&D Organizations	40
4. Evaluation of Researchers' Performances	42

Introduction

Japan enacted the “Science and Technology Basic Law” (#130 Law of 1995), which aimed at “establishing a nation based on the creativity of science and technology.” Based on the Law, the first phase—the Science and Technology Basic Plan (by decision of the government of Japan in July 1996)—and the second phase—the Science and Technology Basic Plan (by government decision in March 2001)—were adopted. The second phase, the Basic Plan, deemed it essential to adopt key policies for setting strategic priorities in science and technology (S&T), research and development (R&D) system reforms, along with the internationalization of S&T activities, to achieve sustainable development of the nation and to attain a national system appropriate for Japan’s international status. It aims to resolve social and economic problems and to contribute to the world by creating and utilizing scientific knowledge. The Plan refers to reform of the evaluation system as a major pillar for developing an S&T system that generates outstanding results.

Based on the first phase, the Science and Technology Basic Plan, the government adopted the National Guidelines on the Method of Evaluation for Government-Funded R&D (by decision of the Prime Minister in August 1997). Based on the second phase, the government then adopted some guidelines entitled “National Guidelines for Government-Funded R&D (by decision of the Prime Minister in November 2001, hereinafter referred to as the Former Guidelines). The aim of these Guidelines was to advance reform of the R&D evaluation system and to bring about improvements in the quality, fairness, and transparency of evaluations. This approach

reflects their results in resources (budget, human resource, etc.) allocation, in securing the required resources for evaluation, and in setting up the organizational framework for evaluations. At this point, the Council for Science and Technology Policy, in its follow-up on the Former Guidelines, was able to confirm that reforms of the evaluation system were going ahead, and that the evaluation process had taken firm root. On the other hand, it was found that the progress of reforms was insufficient, and that new issues had surfaced in connection with the implementation of evaluations. The Committee also proposed a future course of improvement that aimed at (1) evaluation that encourages people to take up the challenge of being creative, and inquires into the outcome; (2) reliable evaluations that live up to international standards; and (3) evaluations that are utilized and that bring about change. On this basis, these Guidelines represent a progressive revision the Former Guidelines.

R&D evaluation referred to in these Guidelines has a different scope of subjects from those subject to policy evaluation under the “Law Regarding Policy Evaluation Conducted by Administrative Organizations” (Law #86 of 2001), but is conducted to the shared goal in principle. These Guidelines consider various factors required for policy evaluation, and take into account the unique characteristics of R&D activities. Evaluation activities, based on these Guidelines, should be approached in the same manner as policy evaluation according to the Law. Evaluation on independent administrative research institutes (including public corporations allocating R&D funds) should be conducted in line with evaluation based on the “Law on the General Rules of Independent Administrative Institutions” (Law #103 of 1999), and those on national universities and inter-

university research institutes, in line with the “National University Law”(Law #112 of 2003).

These Guidelines set out basic policies on the evaluation of R&D. It stipulates that, to ensure implementation of R&D in an effective and efficient manner, evaluating/administering organizations should aim at effective and quality evaluations that are suitable for the particular R&D. This is typically accomplished by conducting or administering evaluations that are matched to the unique characteristics and nature of the respective R&D.

These Guidelines applies to evaluating organizations. That is, organizations that conduct and promote R&D*¹⁾, or third-party evaluating organizations on R&D targeted by these Guidelines (third-party evaluating organizations: *²⁾). The ministries shall enact concrete guidelines that set out the evaluation method and other information in line with these Guidelines. Furthermore, R&D organizations and third-party evaluating organizations shall set out clear rules in line with these Guidelines and the guidelines enacted by the ministries, and conduct evaluations appropriate for the unique characteristics and nature of the organizations and subject R&D.

*1) R&D implementation/promotion organizations are envisaged as follows:

- Ministries
- Universities (including national, public, and private universities), inter-university research institutes, independent administrative research institutes (including

public corporations allocating R&D funds), national research institutes, etc.

- *2) Third-party evaluating organizations are envisaged as follows:
- Council for Science and Technology Policy
 - Evaluation Committee for Independent Administrative Organizations, Evaluation Committee for National Universities, National Institution for Academic Degrees

Chapter 1: Basic Concept and Framework of R&D Evaluation

1. Significance of Evaluation

Evaluation is conducted to effectively and efficiently promote outstanding R&D. This includes R&D of internationally high standards, R&D that can contribute to society / economy, and R&D that pioneers new fields. The significance of evaluation is as follows:

Appropriate and fair implementation of evaluation will help create a flexible, competitive, and open R&D environment in which researchers can exercise their creativity fully.

Conducting supportive evaluations will help advance R&D and improve its quality, identify outstanding, innovative and promising R&D and researchers, and encourage researchers to be more highly motivated. It will also make it possible to formulate better policies and measures.

Actively publicizing evaluation results and outstanding R&D will help the government fulfill its accountability to the public and win broad-based understanding and support for the injection of government funds in R&D.

Reflecting evaluation results appropriately in the allocation of the resources, such as budget and human resources, will help prioritize R&D and make it more efficient.

2. Scope of Application of these Guidelines

R&D evaluations subject to these Guidelines are evaluations on (1) R&D measures, (2) R&D themes, (3) R&D organizations, and (4) researcher's performances. Scope of R&D extends to all R&D conducted with government funds. More specifically, subjects are R&D implemented/promoted by R&D implementation/promotion organizations such as ministries, universities (including national, public and private universities), inter-university research institutes ("universities, etc." hereafter), as well as independent administrative research institutes, and national research institutes. Also subject to the Guidelines are government funded R&D conducted by private organizations or public testing and research institutes, or other government funded R&D carried out overseas.

3. Responsibilities of Evaluating Organizations, Evaluators and Researchers

(1) Responsibilities of R&D Implementation/Promotion Organizations

R&D implementation/promotion organizations are responsible for developing specific evaluation mechanisms (developing evaluation guidelines, setting up an evaluation committee, etc.) according to these Guidelines, conducting strict evaluation, and utilizing the results appropriately to help researchers exercise their potential fully. Other responsibilities include actively providing the public with information about evaluation results and their reflection status. In this process, researchers should be encouraged to bring

their abilities into full play by challenging high goals so as to improve the quality and efficiency of R&D. At the same time, thought should be given to the issue of averting the fact that the extra workload involved in evaluations places an excessive burden on the researchers in conducting their R&D. Related ministries should also be fully aware of their responsibility as presiding ministries in ensuring that they conduct evaluation and utilize the results appropriately and responsibly.

(2) Responsibilities of Evaluators

Evaluators should always make efforts to correctly understand the evaluation subject, be aware of their responsibility to conduct impartial, fair and strict evaluation, and maintain an attitude of severely questioning the researchers' responsibility in implementing R&D. At the same time, evaluators must try to identify innovative and promising outstanding researchers and R&D, and provide appropriate advice to improve and enhance such R&D. They must also be aware that their evaluation results will become subject to future scrutiny by future evaluators, and that the final evaluation is given by the public.

(3) Responsibilities of Researchers

It is extremely important for researchers (where the evaluation subjects are R&D measures, these include the implementers of these measures, that is, those being evaluated) conducting government funded R&D to be fully aware of their responsibility to aggressively challenge ambitious R&D themes, to produce R&D

results, and to strive for the results of their R&D to be ultimately returned to the public and society—that is, to the taxpayers. Moreover, in the absence of results, researchers must take a serious view of the required accountability and responsibility to produce results. Researchers must also be fully aware of the importance of evaluation as part of R&D activities, and actively cooperate with evaluation procedures to enable the evaluator to gain a full understanding of the content of their R&D by providing comprehensive and correct explanations and other relevant information. At the same time, they must be fully aware of the important role evaluation plays from the professional point of view, and take an active part in evaluation procedures.

4. Direction of Evaluation System Reform

The second phase—the Science and Technology Basic Plan—pointed out that “reforming R&D evaluation” is a major pillar in developing an R&D system that produces outstanding results. It also stated that the Former Guidelines would be reviewed as necessary based on their implementation status.

At this point, the Council for Science and Technology Policy conducted a follow-up on the Former Guidelines, charted the future course of improvement (as shown below) to further develop R&D evaluation in Japan, and revised the Former Guidelines accordingly.

- 1) Evaluation that encourages researchers to take up the challenge of being creative and inquires into the outcome

As the opinion that evaluations discourage researchers challenge—and on the contrary, make them atrophy—seems to be quite wide-spread, evaluations from now on will not only inquire into the outcome of R&D, but also attach more importance to encouraging researchers to take up the challenge.

2) Reliable evaluations that live up to international standards

In view of the lack of techniques and human resources required for conducting highly reliable evaluations in Japan, we will improve the evaluation system for upgrading the quality of evaluation, evaluation techniques, and evaluators.

3) Evaluations that are made good use of and that bring about change

A thorough effort must be made to make good use of evaluations in the continuation/revision of R&D, in the allocation of R&D funds, and in formulating better policies/measures.

5. Follow-up on the Implementation of These Guidelines

The Council for Science and Technology Policy should follow up the implementation of evaluation in line with these Guidelines to ensure that evaluations are conducted strictly and evaluation results are put to appropriate use. It should report its opinions to the related ministries and propose revisions to these Guidelines as necessary.

Chapter 2: Common Principles in Conducting Evaluation

To implement R&D evaluation appropriately, R&D implementation/promotion organizations and third-party evaluating organizations should define specific details of evaluation clearly, including evaluation subjects, evaluation objectives, evaluator selection, evaluation timing, evaluation methods and utilization, and handling of evaluation results, while they should seek to improve the implementing framework for evaluation. The common principles to follow in implementing evaluations are as described below.

1. Setting of Evaluation Subjects

Those subject to evaluation should be defined clearly and specifically, with the contents notified in advance to those being evaluated.

2. Setting of Evaluation Objectives

Evaluation is an important tool in strategic decision making, but in itself is not the objective. Those involved in conducting evaluations should be fully aware of this. After having defined in advance how to strategically position the evaluation in question within R&D activities and who will be involved in which activities, the objectives of evaluation should be defined clearly and specifically, with the contents notified in advance to those being evaluated.

For example, the objectives of midterm evaluation of R&D themes could be summarized as being “decision making (including cancellation) in reviewing progress made against the R&D project

and project alterations in accordance with changing circumstances, as well as reflection in allocation of funds,” and those of the follow-up evaluation of R&D measures, as “grasping what the impact on industry or society will be if the measures concerned are implemented, problems inherent in their implementation, and utilizing them for formulating new measures.”

3. Selection of Evaluators

Except when intending to implement self-evaluation*¹⁾, to secure thorough fairness of evaluation, it is important to actively utilize external evaluation*²⁾ by evaluators who are not associated with either the evaluating or the evaluated organizations. Even if it is inevitable to implement an internal evaluation*³⁾, the opinion of an external expert should be obtained if at all possible. When necessary, a private organization can also be commissioned to implement a third-party evaluation*⁴⁾. In these cases, involvement of stakeholders in evaluation should be avoided in principle by giving a clear definition of the scope of stakeholders. If the involvement of stakeholders in evaluation cannot be avoided, the reasons for this must be explained, and an effort should be made to ensure that the evaluators of the stakeholders are of high moral standards, and that the transparency of the evaluation is maintained. Evaluators should be selected with consideration to their age, employer, gender, etc. to ensure sufficient evaluation objectivity. Evaluators should be given a clearly defined term of appointment. This provision does not apply if confidentiality must be maintained for national security or other reasons.

In conducting external evaluation or third-party evaluation, evaluators must be external experts*⁵⁾ in principle with sufficient evaluation capacity, e.g., being well-versed in the field of the respective R&D. To raise the quality of evaluation, the participation of experts in evaluation techniques, experts with cross-cutting experience and so on, should be sought as necessary. When evaluating R&D themes of a large scale or high social interest, R&D measures, R&D organizations, etc., external intellectuals*⁶⁾ should be involved so as to incorporate broad-based perspectives on situations surrounding the R&D. The researchers' performances should be evaluated according to the rules set out by the director of the organization the respective researcher belongs to.

Also, people from the industrial sector, the fields of human / social science, or experts in the industrialization and marketing of R&D outcome should be actively involved as evaluators so as to adequately reflect the socioeconomic needs to the evaluation according to the characteristics and objectives of R&D. In addition, participation of overseas researchers should be sought as required for evaluation under the perspectives of international competition/cooperation and international comparison of the level of R&D. It should be deeply recognized that it is quite important to appoint the evaluators who have good quality to identify innovative, promising researchers and R&D and select outstanding researchers having growth potential.

Evaluators must keep evaluation contents thoroughly confidential, so as to avoid any new conflict of interest emerging between researchers.

- *1) Evaluation in which the R&D implementation/promotion organization that implements the R&D subject to evaluation acts itself as an evaluator.
- *2) Evaluation in which the R&D implementation/promotion organization that implements the R&D subject to evaluation acts itself as an evaluating organization and appoints external persons as evaluators.
- *3) Evaluation in which the R&D implementation/promotion organization that implements the R&D subject to evaluation appoints a person from inside the organization as evaluator.
- *4) Evaluation in which an independent organization different from the R&D implementation/promotion organization that conducts the R&D subject to evaluation acts as an evaluating organization.
- *5) Persons who are experts in the field of R&D subject to evaluation and other related fields, and belong neither to the evaluating nor the evaluated organization.
- *6) Persons who are experts in fields other than the R&D subject to evaluation or other intellectuals who belong neither to the evaluating nor the evaluated organization.

4. Setting of Evaluation Timing

Evaluation of R&D measures and R&D themes should, in principle, consist of an ex ante and an ex post evaluation. Ex ante evaluation, in particular, should, wherever possible, be conducted in consultation with external experts and intellectuals before the ministries and agencies make their budget decisions. When the R&D is expected to take more than five years, or has no defined period, the evaluating organization should regularly conduct midterm evaluations at

standard intervals of approximately three years, for example, considering the objective, content, characteristics, scale, and so on of the respective R&D. Evaluators should also be aware that, depending on the characteristics of the R&D, it is inappropriate to conduct evaluations under short-sighted perspectives that demand quick results, especially in areas such as basic research*), where it is difficult to produce outcome within the short term. For R&D themes where outstanding results are expected, and R&D progress is anticipated, it is necessary to conduct evaluation at an appropriate timing prior to the completion of R&D so that the R&D can continue without interruption.

R&D organizations should be evaluated regularly by the evaluating organizations at a standard interval of approx. 3 to 6 years so as to flexibly adapt to changes in the circumstances surrounding the R&D and have the R&D proceed at a brisk pace at all times. The performance of researchers should be evaluated according to the rules set out by the director of the organization the respective researcher belongs to.

For R&D measures and themes, it would also be beneficial to confirm the social and economic effects (outcome) and spin-off effects (impact) produced as direct results of R&D (output), including secondary effects, at a certain time after the completion of evaluation. This is where the need for follow-up evaluations arises. These should be aimed at adequate and timely assessment of the status of evaluation by academic societies, by practical application of R&D, by technical innovations and by the creation of new social values brought about by R&D. Moreover, where large-scale R&D facilities, buildings, etc.

were set up, follow-ups should be conducted regarding their status of operation, activities, etc. At the same time, they should be used to verify whether past evaluations were appropriate, and reflect the findings in any reviews of related R&D systems, while analyzing to what degree R&D results are utilized. Efforts should also be made to further improve the quality of follow-up evaluation and make it a still more firmly established practice, taking into account its implementation status.

*) “Basic Research” in the intent of this Guideline refers to research conducted for purely intellectual pursuits, including research earmarked for future applications.

5. Setting of Evaluation Methods

Evaluating organizations should clearly and specifically define and notify those being evaluated of the evaluation methods (evaluation technique, evaluation perspectives, evaluation items and evaluation criteria, evaluation process, etc.) in advance, according to the evaluation objective and evaluation subject, so as to ensure fairness, reliability, and continuity, and to conduct effective evaluation.

Especially important for those being evaluated is that they are given specific and clear information about matters such as the “adequacy in comparison with advanced R&D in other countries” and the “attainability of goals and availability of means to achieve them,” as the case may be.

(1) Evaluation Technique

There are many evaluation techniques, covering the entire range from investigation/analysis prior to evaluation to the actual evaluation itself. When conducting evaluation, an appropriate investigation/analysis and evaluation technique should be selected in accordance with the evaluation subject, timing, objective, and available information. In view of the wide range of evaluation perspectives, such as the need for the R&D, its effectiveness and validity, the most suitable method of investigation/analysis and evaluation must be selected on a case-by-case basis.

Evaluation, especially that concerned with results, must be quality-oriented to show the level of R&D results, since the key objective of R&D is to ultimately yield high-quality results. Efforts should be made to apply evaluation methods using specific indicators and numerical values to ensure objectivity in evaluation with due consideration to the characteristics of different research fields. For example, numerical indicators on the degree of achieving the clearly pre-defined goal (including goals achieved in mid-project and their achievement periods), the status of patent utilization, and the results of assessing and analyzing the quality of published research papers by an objective technique can be used as reference data in conducting evaluation.

However, it should be noted that there are situations where applying quantitative evaluation proves difficult, e.g., basic research based on the researchers' free and innovative ideas. The solution in this case is to use objective information and data to the extent possible, but combine it with qualitative evaluation.

Another factor in raising the quality of evaluation is to select outstanding evaluators or, where applicable, an appropriate organizational framework for evaluation.

In the future, it will be necessary to implement more and better investigations/analyses preceding evaluation, and systematically collect and analyze objective quantitative data to form the basis for judgment to improve the reliability and quality of evaluation. For the time being, evaluation should be conducted by selecting one of the currently available techniques. For future evaluations, however, work is in progress to develop and improve evaluation techniques that bring about socioeconomic change through the effects and spin-off effects of ex ante evaluations and follow-up evaluations, as well as quantitative or objective evaluation techniques for basic research.

(2) Evaluation Perspectives

Evaluation should be conducted from the perspectives of necessity, efficiency and effectiveness, with consideration to the perspectives of policy evaluation as set out in the “Law Regarding Policy Evaluation Conducted by Administrative Organizations.” Evaluation is conducted in line with the subject R&D’s international standards. In addition, since researchers must conduct R&D with constant and strong interest in its relevance to society, the perspectives of the humanities and social sciences should also be incorporated in full for some R&D projects.

In view of the key points of trying to advance R&D and raising its quality through evaluation, it is important that evaluation does not become more administrative in character than necessary, and that the difficulty of the themes the researchers have challenged is also taken into consideration.

To prevent excessive concentration of research funds to specific researchers and to promote effective R&D activities, it is important to identify the “efforts”*) made by research representatives and other contributors, and use the data when planning or selecting new R&D themes under the competitive funding system.

*) Called the R&D full-time ratio (%). It refers to the percentage of a researcher’s time exclusively spent for the R&D activities concerned against the researcher’s annual working hours.

(3) Evaluation Items and Criteria

R&D should be evaluated according to evaluation items and evaluation criteria that are appropriately set up under the three perspectives of necessity, efficiency, and effectiveness to match the characteristics of the R&D.

Evaluation items relating to “necessity” include, for example, aspects of S&T significance (creativity, innovation, pioneering nature, development potential, etc.), aspects of socioeconomic significance (stimulation and upgrading of industrial and economic activity, improvement of international competitiveness, acquisition/exploitation of intellectual property rights, creation of

social values (relating to health and safety of the people), contribution to securing national interest, contribution to planning and implementation of policies and measures, and adequacy as government funded R&D (compatibility with national and social needs, compatibility with the organization's founding principles and midterm objectives, necessity/urgency of government involvement, adequacy in comparison with leading R&D in other countries, etc.). Evaluation items relating to "efficiency" include aspects of adequacy in organizational planning and implementation, goal achievement management, cost structure and cost efficiency, R&D methods and approach, and many others. Evaluation items related to "effectiveness" include aspects of goal achievement feasibility and available means, the capabilities of researchers and research representatives, degree of goal achievement, contribution to the creation of new knowledge, (anticipated) direct results, (anticipated) effects and spin-off effects, contribution to raising R&D quality, prospects of practical use/industrialization, contribution to administrative policy implementation, human resources development, contribution to establishing an intellectual infrastructure, and other areas.

At the same time, evaluation criteria should be set clearly in advance so that the judgment basis for the various evaluation items set does not become too vague.

(4) Setting a Flexible Evaluation Methods

R&D evaluation should be conducted with flexibility, e.g., setting appropriate evaluation items, criteria and techniques, and

introducing perspectives according to the evaluation objectives, evaluation subjects, evaluation timing, and characteristics of the R&D (basic research, applied research, development, trial research, etc.*). It is also necessary to review evaluation items, criteria, and other items in line with the rapid progress in S&T and major socioeconomic changes.

In particular, basic research, which may lead to the creation of new knowledge, should be evaluated with a focus on creativity, innovation, pioneering nature, development potential, etc. Yet, the results of such R&D may not necessarily manifest in the short term, but may lead to unexpected development after a long period of time. For this reason, they must not be evaluated with uniform and short-sighted perspectives in anticipation of quick results.

Even R&D that appears to have relatively predictable results may still hold diverse aspects (basic research, applied research, development, etc.) and prove difficult to classify in a simple manner. Therefore, the contents of individual R&D must be examined closely to define specific evaluation methods.

The use of appropriate evaluation indicators allowing for the special character of R&D work should also be considered for R&D areas that cannot achieve visible results in the form of theses, patents, or the like, in the short term, or those serving an infrastructural role for other R&D (trial research, etc.).

On the other hand, it may be helpful in some cases, depending on the nature and progress stage of the R&D, to conduct evaluation

with the focus on the suitability of the organizational and management system of the R&D, the adequacy of the approach to goal achievement, etc.

*) Relatively typical and continuous work, such as various observation studies, collection / utilization of genetic resources, maintenance of measuring standards, trial studies on safety, etc., and dissemination certain technologies.

(5) Averting Excessive Burdens in Evaluation

Evaluating organizations must avert placing an excessive workload on the researchers and having them spend too much time and effort in conducting their R&D. For example, if several evaluating organizations with different objectives conduct evaluations of the same subject, or if several evaluations targeting different hierarchical levels, such as R&D theme/measure/organization, are conducted at the same time, evaluating organizations should conduct their evaluations while maintaining close liaison among each other and utilizing the results of past evaluations to avoid overlaps.

Evaluating organizations must conduct evaluation efficiently according to methods they judge to be appropriate for the evaluation objectives, subjects and timing, and simplify the procedure as much as possible. For example, the evaluation method adopted for large-scale R&D projects should be different from that used for short-term / low-budget R&D themes.

When an evaluation method is simplified or changed, the evaluating organization must indicate reasons for the change, cite criteria and provide an outline of the simplified method.

To prevent that evaluation becomes an end in itself; it will generally be beneficial for R&D implementation / promotion organizations and third-party evaluation organizations to try to bring about consensus among those concerned, and fill the posts of person in charge of evaluation and evaluators with well experienced people. In addition, efforts should be made to develop and use evaluation techniques that reduce the workload inherent in evaluation while maintaining its quality. R&D implementation / promotion organizations should also conduct checks on their R&D projects in advance and prepare related materials as a way of effectively and efficiently utilizing external evaluation and third-party evaluation. Evaluators, for their parts, should also make efforts to use these materials.

6. Utilization and Handling of Evaluation Results

(1) Utilization of Evaluation Results

To show clearly that evaluation helps with strategic decision making, such as in the management cycle*), it is necessary to utilize the evaluation results correctly, in line with previously clearly defined evaluation objectives and utilization methods. In the evaluation of R&D measures, R&D themes and R&D organizations, R&D implementation / promotion organizations must reflect the results obtained by the evaluating organizations in

the allocation of resources (budget, human resources, etc.) depending on the characteristics of the R&D. They should also use them as an incentive to better promote R&D by giving advice on how to improve the quality of R&D, and monitor and disclose the actions taken to fulfill their accountability to the public. Specific examples of how evaluation results can be used are given below, arranged by evaluation timing.

- Ex ante evaluation: Adoption/non-adoption or plan revision, establishment of an outstanding R&D organizational setup;
- Midterm evaluation: Progress inspection and goal management, continuation, discontinuation, change of direction, administrative improvements, improving R&D quality, enhancing the motivation of researchers;
- Ex post evaluation: Checking for achievement/non-achievement of planned objectives/goals, defining the responsibilities of researchers or research representatives, accountability to the public, development of a results database and application to subsequent evaluations, planning and implementation of next R&D stage, formulation of subsequent policies and measures.
- Follow-up evaluation: Confirming effects and spin-off effects, accountability to the public, use to formulate subsequent policies and measures (including review of the objectives of policies and measures).

In addition, the evaluation results relating to the researchers' performances should be reflected in the treatment of the respective researchers.

Specific examples of uses would include use in the screening process for promotion or appointment to a post, reflection in the hard worker's allowance or annual salary, additional allocation of research funds, granting the privilege to conduct R&D in a free environment, continuation of R&D or implementation of the next R&D stage, etc.

- *) One technique to bring about operational improvements is to implement a typical management cycle in the order of plan, do, check, action, and link the final improvements (action) with the next project (plan).

(2) Disclosing Evaluation Contents to Evaluated Parties

After an evaluation, the evaluating organization must disclose evaluation results (including reasons) upon request from the evaluated organization. Determining the evaluation results is the responsibility of the evaluating organization, provided the evaluated organization makes an effort to furnish explanations and information and the evaluator, on the other hand, makes an effort to understand those evaluated. The evaluation results, therefore, must be seriously taken into account. In this connection, a mechanism allowing the evaluated organizations to receive explanations and give opinions about the evaluation results is required. If the evaluated subject finds it difficult to agree with the evaluation results, it should be given the option to voice its objections, supported by sufficient evidence, to the evaluating organization. Evaluation of researchers' performances and other

measurements should be conducted according to the rules set by the director of the organization the subject researcher belongs to.

(3) Publishing R&D Evaluations

Releasing R&D outcomes and evaluation results to the public in the form of an evaluation report will meet the condition of accountability to the public regarding R&D and R&D outcome, ensure fairness and transparency of R&D evaluation, and have the R&D outcomes and evaluation results widely utilized in society and industry.

Evaluating organizations must actively publish R&D outcomes and evaluation results in an easy-to-understand form using the Internet and other means, and reflect public opinion in the evaluations as required, taking into consideration aspects such as the protection of personal / confidential corporate data, national security, and intellectual property rights.

The results of evaluation of researchers' performances and of the screening process for the competitive funding system, among other measures, must be handled with care under the perspective of maintaining privacy on personal data, preserving the integrity of intellectual property, and other issues.

The evaluation report should at a minimum incorporate the elements listed below in a uniform and easy-to-understand form.

- About the evaluation subject: R&D title, implementer, outline, budget, etc.
- About the evaluation objective: clear and specific objective in utilizing evaluation results.
- About the evaluating organization: names of evaluators, philosophy behind evaluator selection.
- About R&D outcome: R&D outcome, other effects or spin-off effects.
- About evaluation results: evaluation method (evaluation technique, perspectives, items, criteria, process, etc.), evaluator's opinion, evaluation conclusion.

The names of evaluators should also be published within a certain period of time after the evaluation, so as to ensure their accountability for the evaluation results. In the evaluation of R&D themes covered by the competitive funding system, special consideration should be given to preventing concerned parties from identifying which evaluators are in charge of which R&D, so as to avert conflict of interests arising between researchers.

7. Administering an Effective and Efficient Evaluation System

R&D evaluations vary widely with regard to implementation organization, evaluation subject, evaluation timing, and so on. Government-funded R&D, in particular, is a multi-strata process implemented in the hierarchical structure between different organizations or within organizations, and in a chronologically interrelated and continuous manner. This process in its entirety must, therefore, be administered in an effective and efficient manner.

To this end, R&D implementation/promotion organizations must make an effort to improve the functions and effects of the evaluation system as a whole by appointing a person responsible for running the evaluation system, and by setting up the infrastructure and organizational framework for evaluation and linkage/utilization of evaluations within the responsibility of each organization.

(1) Administering Evaluations in a Hierarchical Structure

R&D, from the viewpoint of the bodies which are implementing/promoting it, takes place within a hierarchical structure encompassing the Council for Science and Technology Policy—which has a commanding view of S&T policies as a whole—the ministries which implement S&T policies dividing them into different fields, and the independent administrative research institutes that implement R&D activities under the ministries' jurisdiction.

From the angle of the R&D subject to evaluation, we often find the kind of hierarchical structure where, for example, the activities of a certain R&D organization cover a range from major policies (or basic policies) to a specific system, the various programs under this system, and the individual themes adopted under it.

In such cases, evaluation often is implemented as a multistrata process at each hierarchical level. But because multistrata evaluation easily turns problematic due to duplication, an effort should be made to clearly define the range of responsibility of

individual evaluations and create an organic linkage between evaluations and their utilization.

In evaluating a system, for example, it is more efficient to use results of an evaluation implemented program-by-program for system evaluation, instead of checking every little detail down to individual themes in the programs covered by the system. This lends uniformity and consistency to the evaluation at each level, and imposes the need for responsible implementation. Another approach to implementing evaluation efficiently and effectively at each hierarchical level is to appropriately combine self-evaluation or internal evaluation, external evaluation, etc. or consolidate the information required for evaluation in a form suitable for common use.

Through this kind of linkage/utilization of evaluation that proceeds level-by-level and that takes the characteristics of the specific organization or R&D into account, an efficient evaluation system, reaching from individual themes to higher level organizations and measures and policies, can be created.

(2) Administering Chronological Evaluations

In the same way as R&D measures, R&D themes, etc., are evaluated within their life cycle by ex ante, midterm, ex post and follow-up evaluations, chronologically they are often subjected to serial evaluations.

In this case, the evaluations at each hierarchical level should not be implemented in a piecemeal fashion; efforts should be made to give the evaluations continuity and consistency by utilizing the information from earlier evaluations for later ones, or by checking and other means, and by maintaining organic linkage chronologically.

For example, to pursue the flexibility of the evaluation system and consistency of evaluation, new evaluators should be added, while others are kept on throughout a series of evaluations from ex ante to follow-up evaluation, depending on the characteristics of the organization and R&D.

It is also necessary to administer evaluation by collectively managing the respective series of evaluations to be able to backtrack the respective R&D process, and to utilize the results of ex post and follow-up evaluations for improvement of R&D themes and measures in the next stage. Through this kind of evaluation administration, an evaluation system that gradually improves the quality of individual R&D and related higher level measures must be established.

(3) Review of Evaluation System

R&D implementation/promotion organizations should review and revise, as necessary, the administration and functionality of the evaluation system at intervals appropriate for the organization and R&D concerned.

Necessary aspects to be included in the review would be, for example, whether the evaluations at each hierarchical level are conducted properly in line with the guidelines, whether evaluations are not futile or mere formal, whether the role allocation between the evaluators involved in the evaluation is clear and appropriate, whether there is sufficient communication between evaluating organization, evaluators and evaluated subject, etc.

8. Improvement of Evaluation Implementation System

The systems for implementing evaluation must be improved to secure the issues discussed in items 1 to 7.

R&D implementation/promotion organizations must define and disclose a specific mechanism of implementing evaluation to conduct evaluation of high quality and effectiveness according to R&D characteristics. To develop an evaluation system of the highest international standards, it is also necessary to conduct evaluations and investigations/analyses required for these, to secure the budget required for evaluation and related investigations/analyses, to set up an evaluation system, and to develop human resources for conducting high-quality evaluations, etc. In this event, it should be considered to partially allocate research funds to evaluation duties as required.

(1) Developing and Securing Human Resources for Evaluation and Raising the Level of Evaluations

R&D implementation/promotion organizations should set up an evaluation department, appoint a person in charge of evaluation, and assign personnel with research experience, including younger people, from within or outside Japan, depending on their abilities, so as to establish a framework for effective and efficient administration and upgrading of the evaluation system. Also, organizations allocating competitive R&D funds should develop a management system with veteran researchers appointed as full-time program directors*1) and program officers*2) to adequately administer the competitive R&D funding system, to control the process of R&D theme evaluation, to conduct high quality evaluation, to support outstanding research, and to improve the quality of R&D themes applying for evaluation.

In addition, efforts should be made to proliferate evaluation techniques through training, symposiums, and the like, to implement research surveys aimed at upgrading the evaluation system, to train/utilize third-party evaluation organizations, to consider a system of gathering human resources with evaluation expertise in the evaluation department, and to consider offering incentives that would make participation in an evaluation beneficial to the evaluator's social and personal standing. Efforts should also go toward developing a mechanism to evaluate evaluators and setting up other evaluation support systems to establish a framework for developing human resources for evaluation, including younger people (experts in specific fields, well versed in evaluation and personnel of R&D implementation/promotion organizations, researchers specialized in evaluation, etc.) and improving evaluation performance.

At the same time, human resources capable of identifying innovative, outstanding researchers and R&D to work as evaluators and evaluation staff must be developed and secured.

- *1) A highly ranked person with research experience administering the competitive R&D funding system and its operation.
- *2) A person with research experience in charge of selecting individual programs and R&D themes under the various systems, as well as their evaluation, follow-up, etc.

(2) Database Development and Introduction of Electronic Systems for Efficient Evaluation

To streamline evaluation duties (evaluator selection, evaluator evaluation, etc.), to avert unnecessary duplication of R&D, and to promote effective and efficient R&D development, ministries and related R&D organizations must develop and administer a database to be input by the researchers themselves that covers research objectives, R&D fields, researchers (including their “Effort” data), funds (system, amount), R&D outcomes (thesis, patents, etc.), evaluators and evaluation results (evaluation opinions, etc.), sorted by R&D theme. Furthermore, to enable cross-ministerial use of available information, the Council for Science and Technology Policy should take the initiative in developing a system that allows real-time continuous monitoring of the progress of government funded R&D. The Cabinet Office should take the

lead in readying this system for common use by ministries concerned.

In addition, an electronic system must be introduced for higher efficiency in areas such as application processing, document-based screening, and disclosure of evaluation results.

Chapter 3: Facts to Pay Attention to, with Different Evaluation Subjects

Evaluations should be implemented according to the common principles described in the previous chapter, and target “R&D measures,” “R&D themes,” “R&D organizations, etc.” and “researchers’ performance” for each evaluation subject, paying attention to the points outlined below.

1. Evaluation of R&D Measures

R&D implementation/promotion organizations and third-party evaluation organizations should pay special attention in their evaluations to assessing whether R&D measures, such as strategies, systems, and programs enacted to achieve the objectives of governmental and ministerial policies, and the founding objectives of organizations, are appropriate in light of government policies. They must also determine whether they are effectively and efficiently implemented in coordination with related policies, and whether the outcome is adequate in light of the objectives of these measures (or whether such outcome is anticipated).

The results of evaluating R&D measures must be reflected in the review and improvement of the relevant R&D measures concerned to allow for improved measures in the future.

Further effort will be required to bring about improvements in the evaluation of R&D measures and getting it better established, taking its implementation status into account.

2. Evaluation of R&D Themes

An R&D theme is a separate topic on which the researcher conducts specific R&D, its objective, nature (basic research, applied research, developmental research, etc.) and field being extensive and diverse. Therefore, the evaluation technique, evaluation items, etc. must be appropriate for the theme's objective, nature, field, etc.

R&D themes are classified into “competitive fund themes,” which are competitively selected and implemented from public submissions, “priority fund themes”—which are promoted with priority according to clear objectives / goals defined by the government—and “base fund themes,” which are conducted with funds constantly distributed to R&D organizations. Each of these types of themes must be evaluated, taking items (1) to (3) below into account.

As for government-funded R&D themes conducted by private organizations or public research institutes (commissioned parties, joint research partners, etc.), evaluating organizations must also consider evaluation methods for private organizations or public research institutes, and implement these evaluations adequately based on common evaluation principles.

(1) Competitive Fund Themes

R&D themes financed with competitive research funds can be divided into “basic research based on researchers' innovative

concepts” and “research with designated objectives” for achieving specific policy purposes.

“Basic research based on researchers’ innovative concepts” should be made subject to peer review by highly talented experts in light of the related international standards. “Research with designated objectives” should be subject to evaluation that emphasizes not only the S&T perspective, but the socioeconomic perspective as well.

In the ex ante evaluation, which is a screening process as to whether to adopt a particular R&D theme, it is important to respect minority opinions and avoid overlooking novel ideas and creativity. Also, researchers with no or little previous record of R&D submissions (emerging researchers, researchers from the industrial sector) should be evaluated appropriately with the focus on research contents / plans to grant R&D opportunities.

In the case of group research, evaluation should also be made on role allocation among participating researchers, implementation systems, and systems of accountability (including the responsibility of representative researchers).

For themes that are anticipated to yield outstanding results or lead to R&D development, an adequate mechanism for setting a new R&D schedule before the end of the period in progress should be adopted to allow evaluation to be implemented in a way that ensures continued R&D without interruption when the first R&D period ends.

(2) Priority Fund Themes

R&D themes with priority funds should be evaluated as to whether the project matches the purpose of the higher-level R&D policies, whether it has been chosen appropriately, and whether the aimed at specific outcome is (or is anticipated to be) achieved. It should also be noted that the progress of S&T and changes in the socioeconomic situation affect and alter evaluation items, criteria, and other factors. Evaluation from socioeconomic perspectives should be emphasized, particularly in applied research, developmental research, and other related fields.

For large-scale projects, strict evaluation must be made especially on the system of accountability (including the responsibility of representative researchers), and on cost-effectiveness. External evaluation or, as required, third-party evaluation should be encouraged for large-scale projects so as to ensure evaluation objectivity and fairness. In addition, to ensure public understanding at the ex ante evaluation stage, for example, the R&D details, and at the ex post evaluation stage, the R&D outcomes, should be publicized widely over the Internet and through other means, with public opinion being incorporated into evaluation as necessary.

The evaluation for international joint projects must examine international role allocation, international contribution, and the significance / benefits in terms of national interest.

(3) Base Fund Themes

R&D organizations must clearly state the founding objectives of their organizations, and define adequate rules for evaluating R&D, reflecting the evaluation and its results in resource allocation under the responsibility of their directors. In the actual process, they should efficiently and appropriately utilize evaluation made through citation by researchers in related fields.

3. Evaluation of R&D Organizations

R&D organizations should be evaluated according to their founding objectives, medium-term goals in terms of their implementation / promotion of organizational administration and R&D. Third-party evaluation should be actively used to enhance evaluation objectivity and fairness.

As for organizational administration, R&D organizations should be evaluated on administrative actions taken to achieve research objectives / goals or to improve the R&D environment, incorporating the perspective of efficiency. Evaluation items for organizational administration include, for example, improving the support system / intellectual infrastructure, developing / securing human resources, promoting the fluidity of human resources, promoting industry-academia-government collaboration, and conceivable social contributions in their areas of expertise. Evaluation items must be chosen appropriately according to the founding objectives and medium-term goals of each R&D organization.

In the area of R&D implementation / promotion, R&D organizations should be evaluated in the overall performance of the R&D measures/themes they implemented / promoted. In this case, the results obtained from evaluation of other measures and themes should be used as appropriate. Evaluation results should be reflected in the allocation of resources, such as the administrative budget and human resources.

R&D organizations are administered under the discretion of directors, and evaluation results therefore lead to the evaluation of directors' performances.

In addition to the above, evaluation should be conducted as described below according to the character of each R&D organization.

Universities must conduct self-inspection / evaluation strictly as defined in the School Education Law. In doing so, attention must be paid to various characteristics, such as respect to university independence and combined promotion of education / research. As for national university corporations and inter-university research institute corporations, national university corporations should be evaluated for their interim performance (the extent of achieving their medium-term goals, and other measures) under the National University Corporation Law by the Evaluation Committee for National University Corporations (the status of educational development is evaluated by the National Institution for Academic Degrees, and the result must be respected), and the Ministry of Education, Culture, Sports, Science and Technology will adequately reflect the evaluation results in the allotment of operation subsidies.

Third-party evaluation by the National Institution for Academic Degrees, for example, should be promoted on items such as education, research, social contribution, and organizational administration.

Independent administrative research institutes, on the other hand, should be evaluated by Administrative Evaluation Bureaus of the applicable ministries on their interim performance (extent of achieving the medium-term goal, etc.) under the Law on the General Rules of Independent Administrative Institutions. The ministries will adequately reflect the results in the allocation of operation subsidies, while independent administrative research institutes should also make efforts to incorporate the findings into their organizational administration.

Regarding private organizations and public research institutes that conduct R&D (commissioned or joint research) with government funds, evaluating organizations must examine their R&D system administration for applicable themes within the necessary range, from the perspective of ensuring effective and efficient application of government funds.

4. Evaluation of Researchers' Performances

The directors of R&D organizations are responsible for defining and implementing appropriate and efficient rules for evaluating researchers' performances, according to their founding objectives. In doing so, they must evaluate quality and focus on not only R&D but also on the activities relating to R&D, such as R&D planning / administration, evaluation activities, social contributions, etc., giving

full consideration to the fact that researchers have diverse abilities and aptitudes. In the case of universities, attention should be given to their dual functions of research and education.

At the same time, consideration should be given to the difficulty of the R&D themes challenged by the researchers and similar factors, while taking an approach that encourages researchers to resolutely take up the challenge, rather than causing them to atrophy.

In addition, cooperation from research supporters is essential for promoting R&D. To this end, we must appropriately evaluate their expertise, level of contribution to R&D promotion, and other factors.