Abstract

Under the belief-system of bureaucratic infallibility which is epitomized by the “Duality of government” (J. Iio), a fall of trust in the bureaucracy which was caused by a series of maladministrations at the 1990s is a serious challenge to the existing regime. Seen in the context of “latent function” (R. K. Merton), the system of policy evaluation has been introduced to cope with this crisis as it is believed to have a faculty of legitimizing the bureaucracy by means of capitalizing on its renowned professionalism. This paper aims at expounding how, and to what extent, the existing system of evaluation which is characterized as “enforced self-evaluation” has measured up to such an expectation. The exposition that follows is founded basically on (1) the summary review of evaluation reports drawn up by ministries and agencies, (2) the results of a questionnaire survey conducted by Mr. Tanabe, and (3) a detailed analysis of the evaluation report produced by a particular ministry. Though it is too early to draw a conclusion in any definite form, the expectation seems to have been met to some extent, especially through the mechanism of convertibility of policy evaluation and evaluation of organization.
The System of Policy Evaluation in Japan

Daiichi Ito

Introduction

The aim of this essay is to bring up an interpretation of the function of the existing system of policy evaluation. In doing this a framework of analysis — “manifest function and latent function” — is used which was set out originally by R. K. Merton (Social Theory and Social Structure, 1957, Chapter 1) to distinguish intentions of a structure and its consequences. Here, modifying the original framework, the dichotomy of manifest function and latent function is taken to designate the formal, professed function of a system and its informal, real function respectively. When applied to the system of policy evaluation, this dichotomy may mean that, while its manifest function is to literally evaluate the effects of policies as well as to feed its results back to the planning of policies, which may be called as practical — or, in W. Bagehot’s usage, “efficient” — effect, its latent function can be found in an effort to capitalize on the symbolic — or “dignified” — effect of the system in order to enhance the trust in the government, or in the bureaucracy that bears it.

In reality, these two functions have been pursued simultaneously. This is rather as a matter of course since (1), seen in perspective, we can make generalization that, while there is no need for the symbolic effect to be matched off squarely against the practical effect, a divergence between them must be kept within a permissible range. In other words, there must be a certain degree of interacting, or correlating relationship between these two effects. If a divergence goes beyond that range, the system of government will lose its effectiveness and become destabilized. And (2), this condition of interacting relationship, which had been met satisfactorily up to the 1970s, was broken off around 1980. Since then, in spite of the fact that the bureaucrats were confronted with the necessity of innovating their values and skills under the radical change of social and economic conditions epitomized by “globalization” and “a revolution in IT”, they failed to do that. As a result of this, the practical effects of their activities were getting degenerate. Whereas, on the side of the symbolic effect, the bureaucrats seem to have lost themselves in a happy memory of economic “success” in the past, avoiding to face up to the hard truth that their values and skills fell behind the times. As the alienation from “the times” was accelerated, their belief in infallibility of the bureaucracy was swollen to a prejudiced impression. No doubt, this was an unmistakable symptom of the abnormal situation. Trust in the bureaucracy would
be lost in inverse proportion to the swelling of such a fallacious belief. It is just to cope with these difficulties that the system of policy evaluation was introduced. A merit of policy evaluation is said to lie in its faculty to legitimize the government by capitalizing on the professionalism of bureaucracy, without having recourse to the troublesome process of democratic procedure. Therefore it is expected that the system of policy evaluation would restore the interaction between the practical effect and the symbolic effect without running a risk of altering the relative power position of bureaucracy in the political system.

What is worth noting here is the hard fact that, even if policy evaluation is useful for recovering trust, it is also conceivable that it can work against recovery, particularly when the result of evaluation of policies — or of organization as will be discussed in Section 4 — turns out to be unfavorable. In fact, Mr. Tanabe has directed our attention to the possibility that the same system is used as a tool of ruining the reputation of bureaucracy (K. Tanabe, “Policy Evaluation”, in The Foundation of Public Administration (in Japanese), 1998). This means that the problem of whether policy evaluation works for or against the enhancement of trust is a matter of what M. Weber calls “chance” and that the bureaucrats have actually taken it. The revelation that introduction of the system of policy evaluation was initiated from within the bureaucracy itself (cf. Nihon Keizai Shinbun, 14 April 1999) can be seen to attest to the truth of such reasoning. If the existing system of policy evaluation proves to have been effective for recovering trust in the bureaucracy, it would undoubtedly be due to this factor of risk-taking decision.

What is expounded below is just a scenario that would lead to the recovery of trust. Perhaps we must wait and see if this scenario comes true or not.

1. Historical background of policy evaluation
(1) Persistence of the hegemony of bureaucracy

“The Japanese have laws, legislators, a parliament, political parties, labour unions, a prime minister, interest groups and stockholders. But one should not be misled by these familiar labels into hasty conclusions as to how power is exercised in Japan” — this is a warning given by a foreign observer who is well acquainted with the politics of Japan (Karel van Wolferen, The Enigma of Japanese Power, 1989, p.25). In my view, this warning is particularly relevant to those labels that designate the phenomena related to, but called by the name other than, the bureaucracy. To take an example, Mr. J. Iio argues that the label of “parliamentary
cabinet system” which has been provided for in the Constitution that replaced the old one immediately after the World War II is misleading since members of the Cabinet, i.e. ministers, regard themselves, not as agents of the Prime Minister who appoints them on behalf of the parliament, but as agents of the bureaucracy of respective ministries they are appointed to. Taking these conditions into consideration he insists that the label of parliamentary cabinet system should be renamed as bureaucratic cabinet system (“Duality of Government and Politician – Bureaucrat Relationship in Japan”, Leviathan (in Japanese), Vol.34). By the way, in terms of the legal theory, such nullification of the Constitutional provision was made possible by having recourse to the time-honored principle of “Buntan-kanri-gensoku” – translated freely as “divided sovereignty” – which is stipulated confirmatively in the Cabinet Law. This implies that each ministry should be seen as a sovereign body within its jurisdiction, ruled exclusively by its own bureaucracy, to negating the Prime Minister’s leadership on behalf of the Parliament. As has been noticed by students of comparative administration, in Japan, a ministry is to be established by its own establishing law, which provides for the internal structure and management as well as the distribution of ministry’s duties and authorities among organizational units. Actually, this establishing law is a successor of the Imperial Ordinance on the establishment of authorities under the pre-war Tenno-regime and, as such, lays the groundwork for “divided sovereignty”. Thus, the whole of establishing laws may be seen as its title deed.

Then, how does the gap between the label and the substance come about? The major reason for this may be found in an unfortunate incident that, while Japan, from the nature of its politics, cannot be governed effectively without the bureaucracy, the new Constitution fails to give an appropriate position to it. (The old Constitution gave such a position to the bureaucracy by defining the “Tenno” properly, since the “Tenno” was the word that indicated both “Tenno” as a ruler and the whole of the bureaucrats who were regarded as his servants). By chance, this hard fact that the bureaucracy is indispensable has been expounded lately by Mr. W. Omori as persistence of the system of “Kan”, i.e. the system in which the autonomous group of elite bureaucrats dominates the formulation and implementation of policies (The System of “Kan” (in Japanese), 2006). According to him, the current system of “Kan” is a survivor from the old “Tenno-regime”. This implies that the postwar reform that brought about the new Constitution was no more than “succession” in the ecological sense of the word, which means that only the label, not the substance, has changed as far as the elite bureaucracy is
concerned. This point can best be exemplified by the failure of the Civil Service Law, enacted immediately after the War (1947), to transform the bureaucracy that was elitist and highly autonomous to the one that is obedient and responsible to the Parliamentary Cabinet. Namely, it was originally intended to replace the closed system of “Kan” with the open system of personnel management, i.e. “job-classification system”. But, confronted with the fierce opposition of the elite bureaucrats, an amendment to the Law was made immediately after its enactment, to the effect that the enforcement of the provision for “job-classification system” should be suspended “for the time being”. This amendment has not been withdrawn yet, which implies that the “for the time being” clause continues to be valid for more than half a century. Such a state of affairs may be seen to confirm the indispensability of “Kan” – the closed circle of elite bureaucrats – in Japan (ibid., pp. 22-49).

What, then, is the essential feature of the system of “Kan” in the current context of policy evaluation? Though the idea of policy evaluation was first discussed in a working group of the (First) Provisional Council for Administrative Reform (1961-4), an effort to put it into practice turned out to be abortive. It was in the course of investigation and deliberation at the Administrative Reform Council of Prime Minister Hashimoto (1996-7) that the introduction of policy evaluation was discussed seriously, which was eventually brought to fruition in the form of Policy Evaluation Law (2001). In its final report, the Council has made an important statement that “in times of the rapid change in the circumstances, it would lead to shrink, postponement and delay of policy judgment if, among all the other actors being regarded as fallible, the infallibility is required of the bureaucracy alone. On this occasion, it is necessary, through turning the way of thinking and presupposing that the government also fails, to provide for a condition in which policy evaluation and change are continuously encouraged.” In all likelihood, such infallibility of the bureaucracy – the trust in it – is the very essence of the system of “Kan”.

It is true that the bureaucracy is invested more or less with trust in many countries. What is peculiar to Japan is the fact that, among many political actors, only the bureaucracy is presumed to be infallible and, therefore, invested with trust exclusively. Then, how has such exclusive trust been secured for the bureaucracy? It is likely that this is due in part to the historical factor, and in part to the functional factor. Historical – it is inherited from the tradition of the “Tenno-regime”, and the mechanism of belief being generated is almost the same as that of what M. Weber called the “traditional” type of legitimacy. Under the prewar system of
“Tenno-regime”, his servants, i.e. bureaucrats, were believed to be infallible because their master, i.e. “Tenno”, was believed to be so. Such property of bureaucrats has been inherited to their successors after the war, on the condition that they are all men of unsullied integrity just as their predecessors used to be. Of course, as “Tenno” was deprived of his political significance in the postwar period, the legitimacy vested in him before the war has been broken up among ministries, to which bureaucrats are willing to dedicate their loyalty. This is the very foundation of “divided sovereignty”. Still, there is no change to the mechanism of trust being devoted in “carte blanche” to what is “traditional”.

And functional — characterizing the administrative system for economic growth as “compartmentalized competition” which means that the market is compartmentalized sector by sector by administrative intervention (called “industrial policy”), Y. Murakami argues that, “while this system was the means to achieve the end of economic growth, it was also supported by the very fact of growth” (The Age of the New Middle Mass (in Japanese), 1984, p.107 and p.146). Seen from the functional aspect, trust in the bureaucracy is the product of accumulated success of policies that are considered to be made and implemented by bureaucrats themselves. Obviously the 50s and the 60s were the period when success was being accumulated. The 70s and the 80s may be seen as the period when trust ceased to be accumulated as the result of policies started fluctuating between success and failure, though its stock was kept intact. How about the 90s? Most likely the stock of trust started to be eroded, since a large number of policies adopted to cope with the difficulties after the burst of bubble economy turned out to be a failure, as if “all the things the bureaucrats had planned backfired on them.” Setting up of the Administrative Reform Council can be seen as a desperate effort to overcome such a difficult situation. Therefore the statement in its final report should be understood as indicating the sense of crisis shared by many about the decay of trust.

(2) Decay of trust in the bureaucracy

Corresponding to the duality of its origin, decay of the trust should be explained dualistically. Historical aspect — even though the belief in infallibility is supported by the immutable principle of “divided sovereignty” at the root, the belief itself is not immutable since it relies in part on the fulfillment of the condition that elite bureaucrats must be all men of unsullied integrity. Therefore, if misconducts that involve them were to be revealed, it will do serious damage to the belief. And
this was what actually occurred in a series of scandals, particularly in “the financial scandal” in 1991. It is reported that a senior official was reminiscing that “in the former days, the people accepted all the Ministry of Finance did as what was done for the good of the state and society. Now, after the financial scandal was revealed, everything has been changed” (quoted in Y. Nishimura, The Cause of the Defeat of the Financial Administration (in Japanese), 1999, pp.81-2).

As to the functional aspect, belief in infallibility is dependent on the past record of the evaluation of policies as success, which was maintained roughly until the 80s as was seen above. Such record of success, however, was discontinued around 1990 when “bubble economy” was burst. This was exemplified typically, again, in the case of financial policies. Then, what was wrong with them? It is most likely that the elite bureaucrats in the financial administration failed to cope squarely with the problem of bank-loans that were not secured – the so-called “bad-loan problem”. As a result of this, the Japanese economy was led to a long depression from the 1990s well into the 2000s. Why did they decide not to tackle with the “bad-loan problem”? According to Mr. Y. Nishimura who happened to be Director General of the Banking Bureau in 1992, a position responsible for making final decision, they were confronted with a difficult choice of whether to tackle with the “bad-loan problem” prior to the problem of “reform of the financial system” or to reverse the order. As a matter of fact, they gave priority to the problem of “reform of the financial system”, procrastinating the “bad-loan problem”, with hindsight that they might have made the wrong choice (Nishimura, op. cit., pp.82-3). To be sure, in the sense that procrastination of the “bad-loan problem” caused more serious consequences on the Japanese economy as a whole than those that would have been incurred if the reform problem were to be deferred, their decision should be evaluated as a failure. In the due course, this incident, together with some other ones, led elite bureaucrats to recognition of the grim realities that even the bureaucracy could be fallible as was stated officially in the Final Report of the Administrative Reform Council (1997).

Why did they make such a bad choice? Mr. Nishimura explained the reason for this, saying that “as we have just arrived at the agreement on the reform plan of financial systems (the essence of which was to remove “the firewall” between the banks and the securities houses) after six years of difficult and painstaking talks at the Advisory Council on Financial Systems, we would not like to miss the rare opportunity to put it into legislation” (Nishimura, op. cit., pp.81-2). Behind this explanation, however, lay their real intention to take advantage of this opportunity
of reform for reinforcement of their own ministry. This point has been elucidated fully by Mr. S. Vogel as “tenacity” of elite bureaucrats —

The Japanese regulatory regime ...... has produced a distinct pattern of a regulatory reform, which I call “strategic reinforcement”. The pattern is “strategic” because the government has not liberalized for the sake of liberalization, but has selectively introduced competition in specific markets while simultaneously ensuring that domestic firms survive and prosper with as little disruption as possible. I use the term “reinforcement” because the government has not retreated from intervention but has reorganized its apparatus for intervention and rearranged its particular mix of interventionist policies by reinforcing critical mechanisms of control (Freer Market, More Rules, 1996, p.207).

The base of such “mechanisms of control” is found in their own ministry — its power and authorities. So, “strategic reinforcement” implies reinforcing the ministry itself, even at the sacrifice of the government as a whole. Needless to say, the interests of ministry do not necessarily contradict with those of the government as a whole. In the period of economic growth when plenty of surplus value was being produced, these two were compatible with each other. Under the condition where growth rate was negative, however, they came to be incompatible, the interests of ministry being pursued at the cost of those of the government as a whole.

But, on reflection, why are elite bureaucrats concerned so much about their own ministry? It is most likely that, as is referred to above, this is because their loyalty that had once been devoted to “Tenno” was transferred to their own ministry as a holder of “divided sovereignty”. In short, in place of “Tenno”, the ministry itself came to be emotionalized. This is the reason why “the departmentalism” of a ministry came to be “the most serious deficiency in Japanese administration” after the war (K. Tsuji, Public Administration in Japan, 1984, pp.6 - 7). What is of importance here would be to realize the truth that such emotionalism was a remote cause that prevented the cool-headed realism from being kept by elite bureaucrats and eventually induced them to make the wrong choice. Probably this is the cost we had to pay for having recourse to an expedient of using the surviving group of elite bureaucrats for the government of “reformed” Japan to tide over the crisis that followed the defeat in the war.

Anyway, it proves to be the case that both in the historical aspect and in the functional aspect trust in the bureaucracy was exposed to the crisis of erosion in the 1990s. It was against these backgrounds that the system of policy evaluation was
introduced. Coincidentally, enlightened by a small number of penetrating scholars, like Mr. K. Tanabe, who was one of the principal architects of the system, an idea that policy evaluation could be an expedient for reinforcing the legitimacy of government by resorting to the professionalism of bureaucracy was gaining support among elite bureaucrats. Therefore it is understandable that the system has been expected, in addition to realizing the official — or “manifest” — function of promoting “the objective and rigorous implementation of policy evaluation and of reflecting the results of this evaluation in the planning and development of policy” (Article 1, Policy Evaluation Act), to realize the unofficial — or “latent” — function of recovering the public’s belief in reliability, if not infallibility, of the bureaucracy so that the system of “Kan” would be won back robustly, which is indispensable to the workings of government in Japan.

2. Outline of the system of policy evaluation

What should be noted at the outset is, first, the fact that policy evaluation that has been introduced to Japan is “self-evaluation”, i.e. the evaluation made by the government organization itself that has made and implemented the policy, as is stipulated in the Policy Evaluation Law that “an administrative organ is to evaluate ...... policies under its jurisdiction of its own accord” (Article 3). In this respect, “evaluation” under the present system is distinguished clearly from the similar activities like inspection or audit that are carried out by the third party or others, though these two are regarded erroneously as the forerunner of evaluation. And, second, the fact that this “self-evaluation” of policies is compulsory by law as far as ex post facto evaluation is concerned (as to ex ante facto evaluation, policies related to R and D, Public Works, and ODA are to be evaluated compulsorily) (Articles 8 and 9). Furthermore, such compulsion is enforced by the legal arrangement that “the Ministry of General Affairs is to carry out evaluation by proxy ...... in case the evaluation of policies cannot be secured by the administrative organ in charge” (Article 12). Of course, this means that the Ministry of General Affairs — concretely its Administrative Evaluation Bureau — is supposed to be responsible for enforcing “self-evaluation” of policies by an administrative organ concerned. Here emerges a distinctive feature that characterizes the Japanese system of policy evaluation clearly, i.e. “enforced self-evaluation” (cf. K. Tanabe, “Actual Working of the System of Policy Evaluation and its Impact”, Leviathan (in Japanese), Vol.38).

Under the system of “enforced self-evaluation”, the determination of which
standpoint to adopt and that of which method to use are left to the discretion of an administrative organ, i.e. ministry, concerned. In fact, some ministries, notably the Ministry of Economy, Trade and Industry, use unique methods of their own. But most of them use a similar — standardized — method. This is because they have determined to choose methods from among a repertory provided in the Standard Guideline (2000) prepared by the Ministry of General Affairs. According to Mr. M. Horie, a former senior official who was involved himself in the preparation work, an outline of this repertory is as follows —

Policy evaluation is classified into three types in view of its purpose, method and so on. They are project evaluation, performance evaluation and comprehensive evaluation.

“Project Evaluation” is designed to evaluate a policy for implementing certain projects or programs, usually prior to the decision on that policy, from various viewpoints including the validity of its objective in light of the needs of the people or society or a higher objective, the necessity of its implementation by an administrative organization in view of the desirable way of administrative involvement, and the prospect of obtaining results or policy effects in comparison with estimated costs for the policy implementation. Project evaluation is expected to contribute to the decision making on whether or not to adopt certain policy, which alternative to choose, and so on. Project evaluation can also be made after certain policy is implemented to verify actual effects of the policy on the basis of evaluation undertaken prior to the policy implementation.

“Performance Evaluation” is designed to evaluate the achievement of objective by measuring and comparing actual achievement with the result- or effect-oriented target level objective which is preset by clarifying the ends and means relationship in the policy structure. Performance evaluation is expected to contribute to continuous policy review and improvement. Actual achievement of target level objective is periodically and continuously measured during the implementation process. And at the end of the implementation period, evaluation is made on the performance during the whole period of implementation, final achievement and so forth. Practically performance evaluation is understood almost the same as “performance measurement”.

“Comprehensive Evaluation” is designed to evaluate a policy in a comprehensive manner by conducting an in-depth analysis of policy effects from various angles, identifying problems with policy and analyzing causes for them.
Comprehensive evaluation is expected to contribute to policy review and improvement by providing various kinds of data and information conducive to solving problems related to the policy in question. Comprehensive evaluation is made after certain length of time from policy decision or adoption. Comprehensive evaluation may be understood almost the same as “program evaluation” (Policy Evaluation in Japan, paper presented at the International Seminar on Government Performance Evaluation, Seoul, 2006, pp.4-5).

This repertory seems to be thought out well. But the problem of whether policy evaluation is carried out successfully or not depends to no small extent on the quality of personnel — how far the officials concerned are equipped with the knowledge and skill necessary for carrying out policy evaluation — as well. On this Mr. Horie seems to be a little sceptical, saying that —

In spite of the fact that officials in the Administrative Evaluation Bureau had been experts of evaluation activities called “inspection” and had high professional ethics, even they had felt uneasy when it was decided to introduce the “new” evaluation system. Very much so were officials in other organizations who were familiar with policies in their charge but not engaged in evaluation activities. The Administrative Evaluation Bureau organized orientation and training programs for officials of the bureau and also of other ministries and agencies. Each ministry or agency also organized such programs. In some organizations strong and effective leaders appeared who found much interests in policy evaluation and worked strenuously with passion for reform. Generally speaking, however, human resources for evaluation activities are still limited and the quality or level of skills of them should be raised. Involvement of outside experts is not only appropriate to ensure objectivity of evaluation but also conducive to infuse expertise. To promote effective policy evaluation activities it is necessary to foster core officials for evaluation activities in ministries and agencies. Professional graduate schools for policy studies recently established in public and private universities teach policy evaluation and are expected to educate future evaluators in government and those familiar with policy evaluation outside the government. Meanwhile, outside experts such as public accountants and policy researchers at think tanks can be recruited as full-time officials for certain period (op. cit., pp.8-9).

Then, what is the merit of the system of “self-evaluation”? In all likelihood the most important merit would be the flexibility in making the setup for evaluation. As can be inferred, thanks to the way in which standpoint and method of evaluation
are determined freely by a ministry concerned, such standpoint and method as are suited to a particular policy or a particular organization will be chosen flexibly so that conformism — and ritualism that accompanies it — in evaluation is avoided and more substantial evaluation will be materialized. On the other hand, however, there can be demerits. Among them, the most serious would be deterioration of the quality of evaluation caused by an artifice to keep up appearances — a gloss of good performance — so that financial resources would be appropriated favorably for a particular policy or to a particular organization that is responsible for it. As Mr. Tanabe suggested, in order to prevent such deceptive practices from taking place, it would be necessary to make an arrangement to transfer the system of evaluation from “self-evaluation”, i.e. internal evaluation, to external one every time when it is found that deceptive practices are going on. External evaluation, i.e. evaluation by the third party or others, or a threat of it can be an effective deterrent to the deception. As to the forms of it, audit by the Board of Audit and the constitutional power of the Diet to investigate government activities are referred to (Tanabe, op. cit., pp.96-7).

What is of importance here is to notice the fact that deciding to transfer the matter to some form of external evaluation should be seen as a part of enforcement work. So, we are now moving from a discussion about “enforced self-evaluation” to the one about “enforced self-evaluation”. It was indicated above that enforcement work was assigned to the Administrative Evaluation Bureau by the law. This is because the Bureau is considered to be suitable for the job as a successor of the former Administrative Inspection Bureau that had a long experience of administrative inspection. To be sure, it has something in common with policy evaluation. In addition, the duties of the Inspection Bureau had been extended from the inspection of administrative operation to the one of project-output. Still, it is a long way from project-output to policy-outcome, and policy-evaluation must be founded definitely on the understanding of policy-outcome. Therefore, I am rather sceptical, as is Mr. Horie, about the present arrangement that assigns the job of enforcement to the Administrative Evaluation Bureau.

In my short experience as a watchdog of evaluation activities of the Ministry of Education, I have come to be of the view that the contribution the Administrative Evaluation Bureau can make seems to lie not so much in doing substantial work — be it evaluating by proxy or deciding to transfer — by itself as in concentrating on procedural work of bucking up and instigating an evaluation office of each ministry to carry out evaluation against the inertia of underestimating outcome prevalent
among bureaucrats. If so, enforcement work should be rearranged to some other — competent — organ. Thus, we are led to an conclusion that the system of “enforced self-evaluation”, though clever and sophisticated as it is in theory, must satisfy some conditions if it is to be successful in practice — namely, in addition to improvement of the skill of the personnel at large in carrying out policy evaluation, there must be an administrative organ capable of managing enforcement work properly.

3. Implementation of policy evaluation
(1) Progress of evaluation activities

Every year the Administrative Evaluation Bureau sums up the reports of policy evaluation submitted by the ministries and agencies in a summary report titled “Results of the examination of policy evaluation implemented by ministries and agencies”. According to this summary report for 2005, the total of 11,103 individual evaluations were carried out in FY 2004. Its breakdown is shown in Table 1.

<table>
<thead>
<tr>
<th>category</th>
<th>type or kind of evaluation</th>
<th>ex-ante evaluation</th>
<th>ex-post evaluation</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>the rest of policies</td>
<td>performance evaluation</td>
<td>492</td>
<td>139</td>
<td>139</td>
</tr>
<tr>
<td></td>
<td>comprehensive evaluation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>project evaluation</td>
<td>299</td>
<td>40</td>
<td>339</td>
</tr>
<tr>
<td></td>
<td>sub-total</td>
<td>299</td>
<td>671</td>
<td>970</td>
</tr>
<tr>
<td>designated fields</td>
<td>R and D</td>
<td>207</td>
<td>304</td>
<td>511</td>
</tr>
<tr>
<td></td>
<td>Public work</td>
<td>4,784</td>
<td>4,797</td>
<td>9,581</td>
</tr>
<tr>
<td></td>
<td>ODA</td>
<td>28</td>
<td>13</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>sub-total</td>
<td>5,019</td>
<td>5,114</td>
<td>10,133</td>
</tr>
<tr>
<td>total</td>
<td></td>
<td>5,318</td>
<td>5,785</td>
<td>11,103</td>
</tr>
</tbody>
</table>

(Tanabe, op. cit., p.98)

A little more than half (5785) is ex-post facto evaluation and the rest (5318) is ex-ante facto evaluation. As can be seen clearly, the policies in the Public work field are of overwhelming majority throughout ex-ante evaluation and ex-post evaluation, which may reflect a distinctive feature of government activities in Japan that
attaches extraordinary importance to the public investment in infrastructure.

Of the table, “designated fields” in its lower column signify the fields designated by the Cabinet Order. It is prescribed that the ex-ante facto evaluation is obligatory for the policies in these fields—concretely Public (construction) work, R and D, and ODA at present. As to “the rest of policies” in the upper column of the table, in which the ex-post facto evaluation predominates, the number of cases for performance evaluation is by far the largest (492), compared to that for comprehensive evaluation (132) and that for project evaluation (40). Probably this is because the performance evaluation has been taken to be a serious effort to start the ex-post facto evaluation of policies—not of mere administrative operation—for the first time in Japan. In other words, it is expected that this would remove the old evils of overemphasizing the “Plan” stage—at the sacrifice of “See” stage—in the management cycle of government (cf. T. Masujima, Standpoints of Administrative Management (in Japanese), 1981, Chapter 2). Whatever the reason for this might be, however, one thing seems clear, namely, among three types of policy evaluation, the performance evaluation is regarded as most suited to the purpose of policy evaluation and, in fact, its present system has been administered with performance evaluation at the core of the whole system. As a consequence, considerable amount of manpower began to be cast into the work of performance evaluation.

Then, how does its result look like? In putting it into practice, one thing that was called for persistently by the Administrative Evaluation Bureau was to specify—and, preferably, quantify—the level of achievement in terms of the objective to be achieved. In order to do this, it is necessary to work out indices that stand for elements or aspects of the objective so that the level of its achievement would be specified—or quantified—directly or indirectly, i.e. by means of combining indices aptly. In addition it was advised, or urged, to set these indices in terms of outcome that indicated the net effect of policies on the economy and society—not of output—as far as possible. Table 2 shows the extent of specification and how far indices are set in terms of outcome for each ministry or agency.
It is apparent that there is a wide gap in this regard among ministries. That is, there are ministries, like National Land and Agriculture, of which the level of achievement of objectives is highly specified and the ratio of outcome-related indices to the whole of them is very high, on the one hand. And, on the other, there are ministries, like Justice and Finance, of which the level of achievement is less specified and the ratio of outcome-related indices is rather low. No doubt this gap reflects the difference in nature of policies pursued by ministries. For instance, while it is easy to specify the level of achievement in the case of ministries that pursue the policies of providing public services, like National Land and Agriculture, it is difficult to do so in the case of ministries that pursue the policies of social management. This means that the problem of whether a particular performance evaluation is successful or not cannot be decided unambiguously, i.e. by applying a simple, quantitative yardstick, as is presumed so by the Administrative Evaluation Bureau (Tanabe, op. cit., pp.98-101).

As to the implications which this gap may have to the workings of the policy evaluation system as a whole, a reasoning will be presented later in Section 5.

(2) Evaluation of policy evaluation

How do the people evaluate the present system of policy evaluation, both on
commonsense dimension and on analytical dimension?

First, the evaluation on the dimension of statistical analysis. In order to understand the cultural background to the evaluation of policies, Mr. Tanabe conducted a questionnaire survey of the organizational culture, putting five questions to the Division or “Ka” – i.e. an aggregation of its members in the administrative culture of Japan. What follows is the results of that survey (cf. Table 3).

[Table 3. Dispersion of replies]

<table>
<thead>
<tr>
<th>Question 1</th>
<th>I think so</th>
<th>I think so a little</th>
<th>cannot say either</th>
<th>I don’t think so much</th>
<th>I don’t think so</th>
<th>N.A.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question 1</td>
<td>32.8</td>
<td>25.6</td>
<td>29.6</td>
<td>8.8</td>
<td>1.6</td>
<td>1.6</td>
</tr>
<tr>
<td>Question 2</td>
<td>4.8</td>
<td>10.4</td>
<td>16.0</td>
<td>37.6</td>
<td>28.8</td>
<td>2.4</td>
</tr>
<tr>
<td>Question 3</td>
<td>10.4</td>
<td>26.4</td>
<td>34.4</td>
<td>21.6</td>
<td>3.2</td>
<td>4.0</td>
</tr>
<tr>
<td>Question 4</td>
<td>21.6</td>
<td>48.0</td>
<td>17.6</td>
<td>8.0</td>
<td>1.6</td>
<td>3.2</td>
</tr>
<tr>
<td>Question 5</td>
<td>12.8</td>
<td>33.6</td>
<td>27.2</td>
<td>17.6</td>
<td>7.2</td>
<td>1.6</td>
</tr>
</tbody>
</table>

(Tanabe, op. cit., pp.101-103)

The first question is: “Are the managing staff of your organization cooperative with the evaluation activities?” To this, more than a half (58%) replied positively. The second question is: “Are you still reluctant to making policy-related information public?” To this, only a minority (15.2%) replied positively. This may be due to the consequence of the system of free access to information which was introduced in 1999. The third question is: “Are your fellow members keen in learning from the results of evaluation?” To this, only a third of respondents replied positively, which may imply that a channel of learning has not been established yet. The fourth question is: “Do you think that policy evaluation provides with an opportunity to reconsider the present state of administrative activities?” To this, more than two thirds of the respondents (69.6%) replied positively. This implies that, potentially, the role played by policy evaluation for reconsidering administrative activities is evaluated positively. The fifth question is: “Are you interested in acquiring a skill in policy evaluation?” To this, nearly a half of the respondents (46.6%) replied positively, which indicates that an interest in improvement of the skill in policy evaluation is rather great.

Against the background of these features in the organizational culture, he
starts to inquire into the workings of policy evaluation itself by raising two additional questions. The first is: “does the work of policy evaluation impose a burden on you?” Needless to say, this question represents a cost side of policy evaluation. The answer to this disperses as follows: (1) “Yes, very much” — 38.4%, (2) “Yes, to some extent” — 40.0%, (3) “cannot say either” — 17.6%, (4) “No, not so much” — 0.8%, (5) “No, not at all” — 1.6%, (6) N.A. — 1.6%. In short, more than three fourths of bureaucrats feel that the work of policy evaluation is burdensome. This implies that the cost of policy evaluation is rather high. The second question is: “do you think that the work of policy evaluation is useful for performing duties?” This represents an effect side of policy evaluation. The answer to this disperses as follows: (1) “Yes, very much” — 6.4%, (2) “Yes, to some extent” — 44.8%, (3) “cannot say either” — 30.4%, (4) “No, not so much” — 12.8%, (5) “No, not at all” — 5.6%. Clearly about a half of bureaucrats feel that the work of policy evaluation is effective for the running of organization (Tanabe, op. cit., pp.101-104).

Then, Mr. Tanabe goes on to try to find out the extent to which these sense of burden and effectiveness of policy evaluation are affected by the features in organizational culture cited above with the help of regression analysis. His findings are as follows: (1) as to the sense of burden, it tends to increase the more cooperative the managing staff of an organization come to be. This seemingly paradoxical relation may be explained away by the reasoning that, the more cooperative the managing staff are, the stronger their expectation of the results of policy evaluation becomes, which may turn to intensification of the sense of burden. (2) As to effectiveness, it has been found to be increased, the keener the bureaucrats concerned are in learning from the results of evaluation, and the more interested they are in acquiring the skill in evaluation. These findings have been summed up as follows — “the correlation coefficient between the sense of burden and the effectiveness of evaluation activities is negative (-0.2639), which suggests the relationship that the less the sense of burden is, the greater the effectiveness of policy evaluation will be. As is indicated by a result of this regression analysis, in order to reduce the sense of burden of evaluation activities, as well as to facilitate feeding back its results effectively to the performing of duties, it is necessary for evaluation activities to be carried out autonomously. In other words, the conditions that hierarchical intervention is eliminated on the vertical dimension, as well as that the result of evaluation is fed back directly to the field, i.e. within the Division, on the horizontal dimension, must be satisfied. Otherwise, policy evaluation will be degenerated to a mere ritual” (Tanabe, op. cit., p.105).
Second, on the commonsense dimension, “several opinions which need special attention” are introduced by Mr. Horie. These are as follows.

(1) Mind-set of officials and other people concerned has changed. They became more logical, analytical and conscious in considering policy objectives, policy structure, relationship between certain policies and programs or projects and so forth. They became more result- or outcome-oriented and more conscious of not only efficiency but also effectiveness, outcome or effects of policies in question.

(2) Results of evaluation are reflected in policy consideration and budget making far more than ever before. A number of policies, programs or projects have been changed or even terminated and resulted in cutting budget. There are frequent references in the Diet to policy evaluation system and specific cases of evaluation. “See” or “Check” function has taken sound root in the management cycle.

(3) On the other hand, there is a critical opinion which points out that reflection of the results of evaluation in policy consideration and budget making is still insufficient. It is also pointed out that cases of full-fledged evaluation are still few while cases of performance evaluation are very large in number. As to the content of specific evaluation reports, it is pointed out that qualitative analysis and description (are) prevalent in not a few cases and that the disclosure of relevant data and information is insufficient for outsiders to verify the case.

(4) As to the use of evaluation resources, it is pointed out that too many evaluation targets or cases of evaluation activities and voluminous workload led to the dispersion of limited evaluation resources. It is wiser to allocate evaluation resources more intensively to important policies and less to minor or less important policies (Horie, op. cit., P.7).

4. The case of regulatory policies — an exception or a rule?

At this point, we are going to deviate from the main subject for a while, to cast a glance at another — least explored — subject. The cases of policy evaluation we went through in the foregoing sections were, after all, those of policies for providing services, not of policies for regulation. This is because almost all of the evaluations carried out under the current system of policy evaluation are actually those of policies for providing services, as was referred to in Section 3. In fact, it seems that regulatory policies have never been considered to be the proper subjects of policy evaluation. Only recently the government — the Bureau of Administrative
Evaluation — realized the necessity of taking up this problem and started requiring ministries to try Regulatory Impact Analysis (RIA) of their own policies, probably urged by the recommendation of OECD (OECD, Regulatory Reform in Japan, 1999).

Why is it that regulatory policies have never been considered to be the proper subjects of policy evaluation? There seem to be two explanations for this — motivational and technical.

Motivationally, incentives to execute policy evaluation are at work in the case of policies for providing services, as the budget allocation for various services can be affected greatly by the results of policy evaluation. Whereas, in the case of policies for regulation, there are no incentives to policy evaluation at work since these policies do not have any substantial implications in budget allocation. Far from that, there could even be disincentives to policy evaluation as it is conceivable that the results of it would be unfavourable to the policies and, therefore, invite criticism of the departments in charge. Considering that, unlike policies for services that appear to be beneficial to the people at large, regulatory policies are generally unpopular among them, such criticism could grow strong to the detriment of the prestige of government. In short, a great risk is involved in daring to execute policy evaluation. Hence, the tendency to get indulged in a comfortable idea that regulatory policies do not lend themselves easily to evaluation.

Technically, the evaluation of regulatory policies seems to be attended with a methodological difficulty in Japan. This stems from the fact that regulatory measures — in particular economic ones — are implemented in a dualistic way, the one in the ordinary way which is the same as they are implemented in many OECD countries, and the other in the non-ordinary way which is peculiar to Japan. A British observer realized this, and made an insightful comment to the effect that, while in Britain regulation is a determinant of the government-industry relations, in Japan it is a consequence of the government-industry relations as well as a determinant of them (Richard Boyd, Government and Industry Relations in Japan: A review of the literature, 1980, Economic and Social Research Council, p.35). This implies that, as long as regulatory measures stay determinants, the ordinary method of evaluating policies that depends on the validity of causal relations between policy measures and the impacts on economy and society that are caused by them, such as cost-benefit analysis, can be applied even in Japan. But in the case when they represent consequences, this ordinary method can no longer be applied since these consequences, i.e. the state of compliance with regulatory measures, must have been caused, not by these measures themselves, but by “something else”.
As a matter of course, there would be no causal relations any longer that support the ordinary method of evaluating policies. Thus, some other method has to be worked out to cope with such a case of regulation. What is that? In order to answer this question, “something else” that caused these consequences has to be elucidated.

In a model case of regulation — particularly of economic regulation — in Japan, regulatory measures take the form of “administrative guidance”, which is defined in the Administrative Procedure Law as “the guidance, recommendation, advice, or other acts by which an administrative organ demand, within the scope of its duties or functions, certain feasance or non-feasance on the part of specified person in order to realize certain administrative aims, and those which do not fall within the purview of dispositions (i.e. legal acts).” As it is not a legal act — and thus lacks legal binding-force — it is essential for administrative guidance to be effective that “voluntary” cooperation or compliance is secured from the other party, i.e. individuals or firms. Then how can this “voluntary” compliance be secured? According to a senior official who is well acquainted with the official way of doing things, “they comply because they know that: (i) the agencies are ready to exercise legal authority if necessary; (ii) the agencies are ready to take away any benefits under the incentive schemes; or (iii) they may have to face any unforeseen consequences under the situation that the agencies are in a position to mobilize whatever means at hand, including the power of licensing under other statutes, to indirectly force them to comply” (H. Tsukamoto, Administrative Guidance; discussion paper for the Conference on Administration which Changes, 1994, Paris, p.5). Obviously, here, the mechanism of what C. Friedrich once called “anticipated reactions” is at work (Constitutional Government and Democracy). More importantly, this mechanism of “anticipated reactions” is not confined to administrative guidance. To the contrary, a mechanism similar to this is at work also in the regulation of “disposition”-type, that is, where regulatory measures are stipulated in the legal provisions so that they could be put into force to the letter. As a matter of fact, the workings of this mechanism have been made possible by the sophisticated practice that takes legal provisions as a sort of last resort, not to be enforced literally. Indeed, the history of administration in Japan can be seen as the one that has made efforts to cultivate such a practice very carefully. Many, if not all, of the policy objectives — such as public security and the control of risks — have been achieved with the help of this sophisticated practice, not by simple enforcement of legal provisions. To take an example from the licensing of public (bus) transportation business, it is reported that, when multiple applications were made
for a single business opportunity, what the Ministry of Transport actually did was to
guide — or instruct — applicants to strive for self-adjustment among themselves. On
reflection, however, to guide applicants to self-adjustment so that multiple
applications are drawn together to one would be nothing else but to evade making a
difficult decision. As Professor Morita remarks rightly, “this implies that the
decision is made virtually in the process of negotiation between, or among,
applicants, while the ministry’s job is just to confirm the decision already arrived at
by applicants themselves” (A. Morita, Government Regulation and Bureaucracy (in

Then, how has such a sophisticated practice been formed and maintained? It is
at this point that the intervention of trade association — “a reliable interlocutor” of
the government department (Ronald Dore, Taking Japan Seriously, 1987) — is
needed. Unlike the trade association in Britain which is of “a minor
information-distributing and lobbying role”, that in Japan is one of the major actors
in the making and implementing of policies. This can be seen in the report of a
survey conducted by the Fair Trade Commission (FTC), the government agency
which is responsible for administering the Antimonopoly Law. According to the
report, the total number of trade associations registered at FTC is 14,863 (as of
1990), among which about 1,000 are nationwide and are, therefore, more influential.
A questionnaire survey has been carried out on these influential associations, which
reveals that, although the affiliation to trade associations is voluntary in principle,
most of the firms of the trade join them actually. As to the merits accruing from the
membership of trade associations, nearly three-fourths of firms mention “getting
easier to obtain information from the government agency”, whereas more than 60%
mention “the increase in trustworthiness” and more than 40% mention “the
relations with the government agency becoming smoother.” This is to confirm the
observation that “Japanese firms belong to their industry associations in the way
that a Japanese farmer hardly conceives of the possibility of not belonging to the
hamlet association in the settlement in which he lives” (Dore, op. cit., p.200).

At the same time, the FTC-survey shed light on the functions performed by
trade associations. According to it, 83.5% of associations are performing the function
of transmitting the government intention or information to the member firms,
67.5% of associations that of conducting market-research on behalf of the
government agency, 64.0% of associations that of organizing regular meeting with
the government agency, 55.3% of associations that of cooperating with the
government agency in determining the official standards as well as in making out a
What should be noted here is the fact that these associations that are cooperative with the government are not the result of natural developments in the government-industry relations. Rather, they are the product of deliberate efforts of economic bureaucrats to foster “a reliable interlocutor”, as can be seen in the shrewd observation made by the senior official cited above. “The trade associations are pervasive in Japanese industry. They were legally formed and assigned key roles in the wartime economy: the affiliation in government-sponsored trade associations was compulsory. They retained some legal status in the rationing economy of the immediate postwar period and then were legally terminated with the end of the import quota system. The innovativeness of MITI (=Ministry of International Trade and Industry) officials lay in their idea of adaptation of this organizational heritage for the effective execution of state intervention, by creating the mechanism in which each trade association operates under strong bureaucratic influence and vice versa (to generate voluntary compliance). It is this existence of trade association that made the job of securing collective agreement (i.e. the product of voluntary compliance) which is a prerequisite of administrative guidance easier for MITI. Because of this benefit as well as others of which effects are felt also by private parties, the government agencies frequently have lent their hands or sometimes worked behind the scene in the creation of trade associations in the newly-emerging domains of business” (Tsukamoto, op. cit., pp.5-6).

If so, then, “something else” that caused “consequences” of regulation may be found in a symbiotic relationship of the government agency and the trade association that corresponds to it. Obviously here is an emphasis on self-regulation based upon the spirit of “voluntary” compliance. And such an emphasis has been recognized by an American specialist in comparative law as one of the key features of administration in Japan. In the context of explaining administrative guidance as a regulatory technique of “considerable attraction”, M. Young argues that “rather than relying upon fallible agency determinations, Japanese administrators have turned to enforced bargaining and negotiation. Administrators reallocate bargaining power between the parties so as to assure serious negotiations, but then distance themselves from the process, thereby allowing the parties themselves to make the difficult determinations” (Michael Young, “Judicial Review of Administrative Guidance”, Columbia Law Review, Vol.84, No.4, May 1984, p.943).

As can be gathered from this argument, an emphasis on self-recognition may eventually lead to the recognition of “the regulation of private persons by private
persons” – usually in the form of regulation of business activities by trade associations – as a quasi-official regulation. In fact, the Administrative Reform Committee in the 1990s, while admitting a vast expanse of business activities being covered by “the regulation of private persons by private persons”, gives the warming that some of these regulations are almost indistinguishable form the regulations of the government and, thus, obscures where the responsibility rests (The White Paper on Easing of Regulations (in Japanese), 2000, p.124).

In short, among the bureaucrats in Japan, a strong inclination is at work for dependence on the self-regulating capacity on the part of private firms. But the extreme case of “regulation of private persons by private persons” aside, even in the case of usual regulation based upon the legal provisions – let alone administrative guidance – where such dependence is confined to a relatively small part of negotiation under the “reallocated bargaining power”, if the enforcement of regulatory measures is left unchecked to private parties, there lurks a danger that the negotiation among parties would become “moralized” (R. Dore), or “obligational” (O. Williamson), excessively, to the detriment of the quality of the agreement arrived at. This is because, as is indicated by the prevalence of collusive practices – such as “DANGO” or prearranged tender practice –, firms and individuals in Japan tend to act on particularistic principles attaching importance to personal ties and obligations so that a Japanese version of public order would be established by means of compounding an agreement within each sector.

No doubt it is in this context that the government intervention through regulatory measures is fervently called for. What is expected of it is not to realize – e.g. by rectifying the market failure – a sectoral harmony as a component element of public order. Rather, it is to prevent this from being realized so that the excessive “moralization” – and subsequent favoritism – would be restrained. To quote from the passage by R. Dore, “precisely because Japan is a society where personal ties and obligations are so strong, it needs much more bureaucratic regulation to overcome personal favoritism and corruption than do we who are much more given to acing on universalistic rather than particularistic principles” (p.201). And, in order to meet this need concretely, it is reported that “there is more careful devising of rules (yes, bureaucratic rules)” (p.200). As a matter of fact, this “bureaucratic regulation” – i.e. the use of regulatory measures, not for building a sectoral harmony, but for preventing it from being built – may take the form of confirmation by the Ministry, or the threat of it, of “the decision already arrived at by applicants themselves” which Morita refers to above. In the nature of things this implies that
“confirmation” is not the mere formalities. Rather, it is expected very often to exercise a substantial function of supervising self-regulation by private parties so that excessive moralization or favoritism could be restrained.

The difficulty with this preventive type of regulation would be that it is not easy at all to identify, let alone measure, the outcome of a regulatory policy which is indicated in terms of its preventive effect. If so, it would not be easy at all too to submit it to the procedure of policy evaluation. For one thing, a causal relationship of policy input and preventive effects, knowledge of which is a prerequisite for the procedure, has hardly been explicated. It is most likely that the difficulty related to this is the final reason why the regulatory policy has been regarded as the one that does not lend itself easily to policy evaluation.

5. Convertibility of policy evaluation and evaluation of organization
(1) Typology of performance evaluation — regularities in a diversity

Returning to the main subject, as is referred to in Section 2, the determination of goals to be achieved as well as of methods to be used to achieve goals is left to the freehand of an organizational unit, “Ka”, concerned, the world of implementation is full of diversity. It is, therefore, difficult to draw the configuration of various types of implementation. Nevertheless, it is not impossible to discern some regularities — particularly in terms of the specification of goals to be achieved — in the way in which policy evaluation has been implemented. To take an example of the performance evaluation, the central core of the methods of policy evaluation, it is essential to (1) determine the goals to be achieved by “sesaku”, i.e. program, and to (2) represent these, as far as possible, in quantitative terms so that performance could be evaluated objectively. This is what the Administrative Evaluation Bureau has required ministries to do. And, in connection with this, it has investigated, in the first phase, (1) how many goals of programs have been set and (2) how many indicators have been used to measure the achievement (= performance) of these goals, ministry by ministry, as of the FY 2005. The result of this is shown in [Table4].
It turns out that, with the exception of the Ministry of National Land and of the Policy Agency, goals of programs have been set on the level corresponding to the duties of an organizational unit “Ka”, i.e. Division, as can be seen in the coincidence between the number of goals and the of divisions. At the same time, it proves to be the case that there is a gap in the ratio between the number of goals and that of indicators, from 1.5 in the case of the Ministry of Agriculture to 8.56 in the case of the Ministry of Public Finance.

In the second phase, the Administrative Evaluation Bureau has investigated (1) how many indicators have been utilized for specifying — in particular quantifying — the goals to be achieved, and (2) of these, how many have been set with reference to outcome, which means the impact of the program on the economy and society. The result of this is shown in [Table 2] cited above. It becomes clear that, while the goals to be achieved have been specified and the indicators have been set in terms of outcome, not of output, in the case of a ministry of public works type like the Ministry of Agriculture and the Ministry of National Land, in the case of a ministry of (social) management type like the Ministry of Justice and the Ministry of Public Finance, the goals to be achieved have been least specified and the setting of indicators in terms of outcome has been avoided as far as possible.

Now, by combining the result obtained in the first phase with the one in the second phase, Mr. Tanabe arrives at the conclusion of “corresponding relation”
which is described as follows: “in the case of an organization that tends to set the levels achievement in the form of quantified goals like the type of a ministry of works, there is an inclination to restrain the number of measuring indicators, whereas in the case of an organization that tries to avoid specification of the levels of achievement like the type of a ministry of management, there is an inclination to set a multiplicity of measuring indicators so that the monitoring of institutions would be carried out from a variety of standpoint” (op. cit., p.101).

This conclusion may seem to be persuasive. It would, however, be more convincing if a third type of ministry is supplemented between two extremes of the ministry of works type and the ministry of management type. By “a third type” is meant an organization in which the gap between the number of goals and the number of measuring indicators is neither so big as the ministry of management nor so small as the ministry of works, as can be seen in the case of the Ministry of Education and of the Ministry of Welfare whose gaps are 7.3 and 6.4 respectively. By chance, these Ministries have in common the duties of providing public services to individuals — education services and welfare services. Therefore, they may be classified jointly as the ministry of (public) services, distinguished from the ministry of works that is supposed to provide infrastructure for the economy and society, not for individuals, on the one hand, and from the ministry of management whose duties are to keep procedural orders, both legal and budgetary — not to provide any substantial services — on the other. Thus, by taking (1) the nature of services provided by ministries and (2) the gap between the number of goals and that of indicators with each ministry into account, three types of ministry have been identified, which may help explain the mechanism of latent function as will be seen in the next subsection.

(2) A case of the ministry of public services (Education) — in pursuit of the latent function

In regard to performance evaluation which is the core of the new system of policy evaluation, it has been stipulated in the Cabinet Resolution that evaluation should be carried out with the unit of policy that is considered “reasonable”. And, in the Standard Guideline which is intended to put the Law into effect it has been assumed that (1) the policy system is composed of three levels — policy, program and project — and (2) “reasonable” unit would be on the level of program (“sesaku” in Japanese). This implies that goals to be achieved through policy activities are to be placed on the level of program, or “sesaku”.
Now, it has been found out that there lies a gap between the number of goals and the number of measuring indicators, and that the size of a gap varies greatly between the ministry of works, e.g. 1:1.5 in the case of the Ministry of Agriculture, and the ministry of public services, e.g. 1:7.3 in the case of the Ministry of Education. This is understandable as there is a difference in the “congeniality” of goals to specification, or quantification, between ministries. On the one hand, there are ministries the goals of which are themselves congenial to specification and, therefore, there is no need to set many indicators for specification, as in the case of the ministry of works. On the other hand, however, there are ministries, like the ministry of public services, the goals of which are not congenial to specification and, therefore, there is a need to set a large number of indicators so that the goals would be specified indirectly by combining these indicators dexterously, or in a roundabout way.

Part of the gap can be explained by this “difference in congeniality”. But not all of it — Why? Based on the statistical data compiled by the Administrative Evaluation Bureau, we have been discussing the matter on the macroscopic level so far. Now, in order to answer the question of why not all of the gap can be explained simply by “difference in congeniality”, we may have to change our analytical perspective and start discussing the matter on the microscopic level based on the information provided by individual evaluation reports produced by individual working units — usually Division or “Ka” — under the system of self-evaluation. Then, we shall come across a spectacle which is quite unexpected.

To take an example of the Ministry of Education — a prototype of the ministry of public services — the total of 42 performance evaluation reports were produced in the FY 2003. As a performance evaluation is carried out concretely with a program, or “sesaku”, as its object of evaluation, this implies that the total of 42 programs have been put into effect within the Ministry of Education. By chance, this figure corresponds roughly with the total number of Divisions, or “Ka”, which may suggest that a program is formed and implemented each by a Division as a program unit. As a matter of fact, however, here lies a serious problem — is this correspondence just a coincidence or something more than that? If it is just a coincidence, there is no need to bother about it. On the contrary, if it is something more than that, we may have to investigate “something” that brings about correspondence.

In order to know which is right, let us examine a particular policy of the Ministry of Education, the goal of which is set to “improve academic ability and nurture broadmindedness.” Actually, this policy is composed of four programs —
program 1, the goal of which is to improve academic ability, program 2, the goal of which is to nurture tenderheartedness and to cope with a problem behavior of school children, program 3, the goal of which is to build trust in school, and program 4, the goal of which is to construct rich and comfortable educational facilities. Among these, we shall concentrate on program 2, as a problem behavior of school children seems to have spread rapidly and attracted a great deal of attention.

A glance at the Evaluation Report of program 2 makes it clear that a lot of indicators, seven in all, for measuring attainment, or performance, of the goal have been prepared. They are (1) the number of schools where the moral education has been provided with the cooperation of local talents, (2) the length of hours devoted to “learning by experience” such as staying in farm-house and experiencing farm work, (3) the ratio of schools that are operating the program of work-experience and the internship program, (4) The number of secondary schools in which “school counselors” are stationed, (5) the number of “schooling support centers” that are set up to cope with the problem of “refusal to go to school”, (6) the diffusion of “support-team” that is expected to be the core of local network to cope with the problem behavior of school children and (7) how many special support programs have been worked out for the handicapped children.

Besides these indicators, a set of “reference indicators” that are of a negative nature has been arranged. They include a rate of incidence of bullying at school and that of “refusal to go to school”. By the way it has been revealed in the experimental project that a rate of incidence of “refusal to go to school” is reduced to 1.2% in the school where a school-counselor is arranged, whereas in the school where a school-counselor is not arranged a rate of incidence stays high at 5.0%.

In point of fact, an awkward problem lies here, i.e. these indicators are not the elements or components of the program 2, since there is no equivalent to this on the dimension of legal or quasi-legal constructs. It would be commonplace to define a program in terms of the legal or quasi-legal constructs, as it cannot be communicated, and therefore cannot be said to be existent at all without these constructs. Therefore, the fact that the program of coping with a problem behavior does not have its own equivalent on the level of rule-constructs simply means that it is not in existence in reality. What appears to be a program — the program 2 — is nothing but a projection of what, in the Ministry of Education officials’ mind, ought to be existent.

Then, what is the true character of the above indicators? On close examination, it turns out that the indicators are actually nothing other than “operationalization”,
i.e. conversion to the operational terms, of the duties of an organizational unit, a Division, that are provided for in the Establishing Law and Order of the Ministry of Education. (For example, Article 37, Clause 1, of the Cabinet Order on the Organization of the Ministry of Education stipulates that the principal duty of the Division for School Children is to plan, as well as to help implement, the guideline for educational and career counseling at the primary and secondary schools. The indicator (4) that shows the increase in the number of secondary schools in which school counselors are stationed may be seen as “operationalization” of the above stipulation under the present condition of educational decay). As such, many of these indicators have virtually been in use to confirm the performance of the Division concerned, not the ones that are newly worked out. What the new system of policy evaluation has actually done is nothing but to open these de facto indicators to the public.

If so, then, these indicators — an aggregate of them — should be seen as a substitute for the void program. In other words, these indicators, when they are aggregated, are to work for substantiating the projection of what is thought to be existent at any cost. And, in fact, this is what happened with the Ministry of Education. Here, once the projection is substantiated, the goal of the program, i.e. coping with the problem behavior, has come to be seen, not as a display, but as binding on all parties concerned. Undoubtedly this implies that policy evaluation can be convertible with the organizational evaluation. This fact is to give an answer to the question posed above: is the correspondence of the number of programs with that of organizational units — Divisions — mere coincidence or something more than that? Obviously, the answer is “something more”, and the substance of this “something more” is nothing but a mechanism of convertibility of policy evaluation and organizational evaluation. Because of this mechanism, the number of evaluation, either of program or of organization, coincides necessarily with that of Division.

Let us remember that we started the discussion with the questioning of whether the system of policy evaluation — as its latent function — could be of any use for recovering the trust in the government, and eventually in the system of elite bureaucracy, that had once been injured. Now, the discussion that followed seems to have demonstrated that, first, there are three types of government organization, i.e. the public works type, the public services type and the management type, and, second, as far as the public services type is concerned, there is — again in M. Weber’s usage — a probability that the policy evaluation is converted functionally to
the organizational evaluation. Thanks to this mechanism of conversion, the policy evaluation has come to be seen as useful for recovering the trust in the government organization — and through this the trust in the system of elite bureaucracy as the government organization is controlled systemically by the elite bureaucracy. In fact, Mr. Omori’s argument has been put forward on the assumption that the trust in the system of elite bureaucracy is in the process of recovery partly because of the reforms, including the setting up of the policy evaluation system. Talks with some elite bureaucrats, as well as a variety of attitude surveys of the people concerned, seem to verify this assumption. If there is a problem that remains unsolved, it would be whether there can be any contradictions between the policy evaluation and the organizational evaluation or not — this is the problem that should be tackled next.
References


OECD, Regulatory Reform in Japan, 1999.


