Effectiveness of University Evaluation: Experience and Lessons Learned from the Trial Evaluations Conducted by the National Institution for Academic Degrees and University Evaluation in Japan

Takahiro Saito, Takayuki Hayashi
National Institution for Academic Degrees and University Evaluation

Abstract

From academic year 2000 to 2003, the National Institution for Academic Degrees and University Evaluation (NIAD-UE) implemented evaluations of selected universities and inter-university research institutes on a trial basis to enhance the educational and research quality of these institutions and to promote their accountability. After the trial evaluation had finished, a meta-evaluation was conducted to clarify the outcomes, including successful aspects and problems, through questionnaire surveys and interviews with the evaluated universities and institutes and with the evaluators. As a result, it was made clear that the process and the result of the NIAD-UE evaluation led to the improvement of universities' educational and research activities. However, many of the persons engaged in making self-report and many of the evaluators pointed out the considerable burden of the evaluation process. The apathy of both the other university staffs and the public about the evaluation also remained a problem for making an effective evaluation system.

1 Background
1.1 Trial University Evaluations since 2000

From academic year 2000 to 2003, the National Institution for Academic Degrees and University Evaluation (NIAD-UE) implemented evaluations of selected universities and inter-university research institutes on a trial basis. The trial university evaluation was made to serve the following two goals.

[1] Multifaceted evaluations were undertaken with respect to the educational, research, and social-service activities of each university; the results were served as feedback to each institution, thus assisting it in the process of improving its educational and research activities.

[2] The detailed descriptions of the education and research conditions at each university were made public in an easily understood form, and then the institution could become more accountable to win broad public understanding and support.

In March 2004, we compiled the results of the third cycle university evaluation and reported them to the evaluated universities and institutions, as well as to the general public.

In the original plan of the Japanese university evaluation system, we were to conduct full-fledged evaluations from the 2004 academic year. However, the evaluation scheme had to be modified because of two major changes in higher education policy: a change of the School Education Law and the corporatization of the national universities.

The new higher education policy in Japan was designed to improve the level of the education, research and other activities in universities. The School Education Law,

newly enacted, decrees that all of the universities, junior colleges, technical colleges and law schools are obligated to receive an evaluation conducted by a quality assurance agency with the authorization of the Ministry of Education, Culture, Sports, Science and Technology (MEXT) in order to improve their level of the education, research and other activities. For universities, junior colleges and technical colleges their overall conditions of their education, research, organizational management, and facilities and equipment must be evaluated at least once every seven years. For professional graduate schools (such as law schools) their overall conditions of their curriculum, faculties and other education and research activities must be evaluated at least once every five years according to the institution's goals and objectives [article 69-3 and 70-10]. On the basis of the law NIAD-UE became one of the authorized quality assurance agencies in this new quality assurance scheme on January 2005 and has been accepting applications for the accreditation of academic year 2005.

In addition to the new Japanese accreditation system, NIAD-UE is assigned to implement performance evaluations concerning the education and research activities at national university corporations and inter-university research institutes. The article 34 of the Law Concerning the General Rules of Independent Administrative Institutions applied *mutatis mutandis* in article 35 of the National University Corporation Law states that when evaluating the performance of each national university corporation or other institution during the medium-term objective period, the National University Evaluation Committee established at the MEXT requests NIAD-UE to implement the evaluations concerning the conditions of education and research and respect the evaluation results. As a consequence it is considered that the judgment of the Committee will affect the budget allocation to the national university corporations and inter-university research institutes.

In this manner, the NIAD-UE is expected to take an important role in the new quality assurance system for higher education in Japan. However, introducing such new system means that NIAD-UE will reconstruct its evaluation system as a whole. To make appropriate and effective evaluation system, NIAD-UE must take advantage of its accumulated experiences and lessons learned through the trial evaluations over the past four years.

1.2 Outline of the Trial Evaluations Conducted by NIAD-UE

The trial evaluation was organized into three themes, namely, University-wide Thematic Evaluation (UwTE), Evaluation of Educational Activities in each academic field (EEA), and Evaluation of Research Activities in each academic field (ERA). The targets of UwTE were activities throughout the university, that is, "education services to public", "liberal education", "cooperation with society in research activities" and "international cooperation and exchange activities." EEA and ERA were carried out in nine disciplines for selected department or research courses in the universities or inter-university research institutes (Table 1). The evaluation were implemented on the basis of request from the university's managerial person, namely, Minister of Education, Culture, Sports, Science and Technology for then national universities and the Governor of the prefecture/area for public universities. The number of the evaluated institutions was more than five hundred throughout the trial sessions in total.

In order to support and promote efforts to enhance the quality of education and research activities in each university and to explain these activities to the general public, NIAD-UE conducted impartial evaluations with high transparency and objectivity. After

the evaluation, NIAD-UE provided the results to each university as feedback, and disseminates the results widely to society in an easily understood form.

Table 1 Targeted institutions for the trial evaluations

Category	Starting	Themes / Fields	Number of Institutions
University -Wide Thematic Evaluation (UwTE)	2000	Education Service to Public	98 National Univ. and 14 IuRIs(*)
		Liberal Education	95 National Univ.
	2001	Cooperation width Society for Research Activities	99 National Univ. and 14 IuRIs
	2002	International Cooperation	97 National Univ., 14 IuRIs and 4 Public Univ.
Evaluation of Educational Activities in each academic field (EEA)	2000	Natural Science	6 National Univ.
		Medicine	6 National Univ.
	2001	Law	6 National Univ.
		Education	6 National Univ.
		Engineering	6 National Univ.
	2002	Humanities	6 National Univ. and 4 Public Univ.
		Economics	6 National Univ. and 2 Public Univ.
		Agriculture	6 National Univ. and 1 Public Univ.
		Comprehensive Science	4 National Univ. and 2 Public Univ.
Evaluation of Research Activities in each academic field (ERA)	2000	Natural Science	5 National Univ. and 1 IuRI.
		Medicine	6 National Univ.
	2001	Law	6 National Univ.
		Education	5 National Univ., and 1 IuRI
		Engineering	6 National Univ.
	2002	Humanities	5 National Univ., 1 IuRI and 3 Public Univ.
		Economics	6 National Univ. and 2 Public Univ.
		Agriculture	6 National Univ. and 1 Public Univ.
		Comprehensive Science	4 National Univ. and 2 Public Univ.

(*) IuRI: Inter-university Research Institute

The procedures of the trial evaluation were shown in figure 1. The NIAD-UE evaluation was characterized by the "goals/objectives-oriented", "highly trained peer reviewers", "transparent", and "evidence-based." In compliance with the "goals/objectives-oriented" and "evidence-based" framework set by the Committee for University Evaluation in NIAD-UE the universities and inter-university research institutes conducted the self-evaluation about the targeted activities. The "highly trained peer reviewers" analyzed the self-evaluation report by focusing to the "evidence", which was considered to be important for the third-party evaluation. In addition to this document analysis site-visits were made for EEA, and ERA for engineering. Regarding to UwTE and ERA hearings were implemented in place of the site-visits. After providing an opportunity for objection or comments to the evaluated institutions, evaluation results were finalized. The evaluation reports were then notified to the

institutions and disclosed and disseminated to the general public. All through those processes the "transparency" was assured by opening the minutes to the public except for the descriptions which would have a harmful influence to the specific institutions.
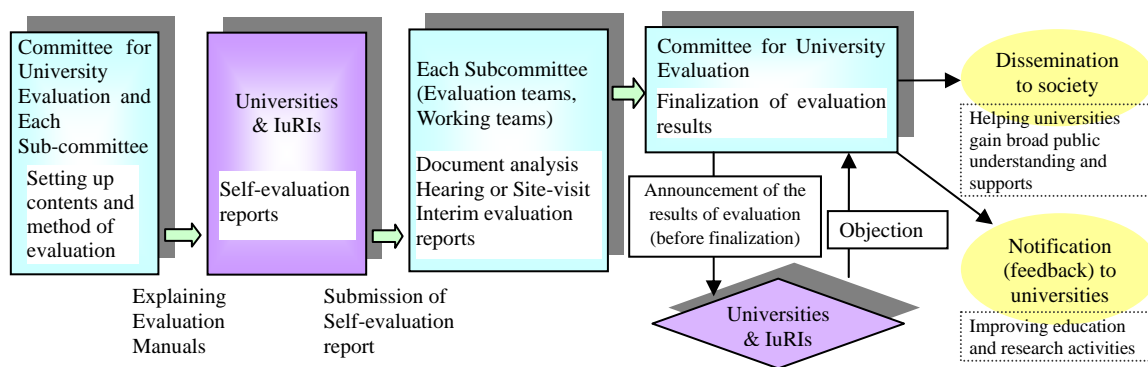
(*) IuRI: Inter-university Research Institute

Figure 1 University Evaluation Process

2 Purpose

In this manner NIAD-UE implemented evaluations of many national and public universities and inter-university research institutes on a trial basis to enhance the educational and research quality of these institutions and to promote their accountability. This "meta-evaluation" study aims at clarifying successful aspects and problems of the trial university evaluation, so that it helps the new evaluation systems to be efficient and effective in order to enhance the quality of education and research at universities and other institutions in Japan.

3 Methodology

All of the factors and activities concerning the trial evaluation were categorized into framework, input, process, output and outcome and all of them had to be verified to ascertain that the each step was appropriate to attain the NIAD-UE missions described above. At the first step a logic diagram of the trial evaluations, which displays the activities and the logical relationships in a chart, was constructed according to the original plan (Figure 2). The verification was conducted based on the diagram.

The following data were collected and used for the analysis.

(1) **Opinion Survey**: Responses to opinion surveys mailed to all of the evaluated universities and inter-university research institutes, and also to the evaluators and concerned organizations (ex. National University Association), from time to time during the trial evaluations were analyzed. The style of the surveys was mainly free responses. All of the collected opinions were sorted into approximately five hundred categories.

(2) **Questionnaire Survey**: Responses to a questionnaire survey mailed to universities and inter-university research institutes in July 2004 were analyzed. The purpose of the questionnaire survey was to figure out the self-evaluation process inside the targeted organization, especially workload and internal system for the evaluation, and the outcomes that appeared inside the

organization as a result of the evaluation, such as reforming their operation, improving the content of their activities, enhancing the performance of their educational and research activities, etc. The respondents were persons in charge of the self-evaluation. The number of responses was 539, and the rate of collection was 98 percent.

(3) **Interview**: In order to perceive the real situation around the evaluation in the target institutions, opinions of the managerial persons and persons who engaged in the self-evaluation were collected. We visited nine universities and one inter-university research institution in September and October 2004 and interviewed them about the evaluation process and outcomes as a complement to the former surveys.

4 Results

In the verification all of the activities were analyzed but we focused into the process, output and outcome in this study because they were more important to develop the new evaluation system.

4.1 Design of the evaluation

(1) Organization structure and quality of the evaluators

The evaluation was implemented by a hierarchical organization composed of the Committee for University Evaluation, Sub-committees by evaluation theme and academic field, and evaluation teams and working teams under the Sub-committees. Some key persons served as the member of multiple organizations to seek stronger ties among these organizations. Therefore, a big problem regarding the organization composition was not pointed out in the opinion survey for the evaluators.

The evaluators must have two days training offered by NIAD-UE in order to fully understand the trial evaluation system. The contents and way of training have been improved year after year such as improvement of the manuals and introduction of simulation. This may be the reason that many positive opinions were received from the evaluators though there were some opinions to request the further training.

(2) Methodology of the trial evaluation

To promote the individuality of the targeted institution, the evaluation was implemented based on the institutional goals and objectives. In the opinion survey for the first cycle this evaluation framework had many complaints that the institution of which the moderate goals and objectives had been set conservatively could get the good result intentionally. However, at the end of the trial evaluation period the answers for the questionnaire survey showed that the goals and objectives based evaluation was accepted as a proper procedure for the progress of the institution's individuality and improvement.

Prior to the execution of the self-evaluation, the Committee for University Evaluation set the hierarchical items for the evaluation by theme and academic field; those were "evaluation items", "elements", and "viewpoints." Though positive responses for those items were obtained at the questionnaire survey, some critical opinions were seen such as neglection of institution's characteristics, fear of making uniformity in the universities, and irrelevancy of those items to the real situation.

For the design based on self-evaluation, there were a lot of opinions from the NIAD-UE evaluator side such as "Information was insufficient only in the

self-evaluation," "The evidence was short to make an accurate evaluation," etc.

As the reasons for the inappropriate self-report, it was pointed out in the opinion survey that the implementation of the self-evaluation in compliance with the legitimate format was not familiar to the universities and a series of explanations and manuals offered by NIAD-UE were too difficult to understand the way to do the self-evaluation. This problem seemed to be decreasing because the contents of explanation at the workshop and a series of manuals for the evaluation offered by NIAD-UE were enriched year after year by accumulating the experiences.

Regarding to the site-visit and hearing, the results of the opinion and questionnaire surveys for the evaluators and self-evaluators indicated that having the occasion of dialog was very effective for the mutual understanding. However, it was also pointed out that the workload for preparing documents and data for the site-visit or hearing was heavy and also the schedule was overcrowded.

4.2 Output of the trial evaluation
(1) Results for the trial evaluation

The evaluation result was described in the report as the stylized terms meaning the extent to which the outcomes or activities in each evaluation item had achieved or contributed to its goals and objectives. As for the way to show the evaluation result the positive opinions were increasing year after year.

Regarding to the appropriateness of the evaluation result itself including the descriptions and judgment of the level of each evaluation item in the evaluation report roughly positive answers were obtained (Table 2). However, the ratio of positive answers to the question, "The evaluation result was commensurable with the restrictions of your university", was relatively low. The more consideration of the characteristics of each university was requested to the evaluation.

Table 2 Appropriateness of the evaluation result

Item	Not agree at all <———> Strongly agree					Total (%)	AV. STD
The evaluation result was commensurable with your university's goals and objectives.	7 (1.3)	54 (10.0)	160 (29.5)	283 (52.2)	38 (7.0)	542 (100%)	3.54 (0.82)
The evaluation result was commensurable with your university's current situation.	10 (1.8)	77 (14.2)	142 (26.1)	271 (49.8)	44 (8.1)	544 (100%)	3.48 (0.90)
The evaluation result was commensurable with the restrictions of your university (environment, resources, characteristics, etc.).	26 (4.9)	132 (24.7)	206 (38.5)	152 (28.4)	19 (3.6)	535 (100%)	3.01 (0.93)
The result of the ratings on the research assessment was appropriate.	3 (4.9)	6 (9.8)	14 (23.0)	33 (54.1)	5 (8.2)	61 (100%)	3.51 (0.95)
The result of the ratings on the social effects based on the research activities was appropriate.	2 (3.3)	5 (8.2)	28 (45.9)	23 (37.7)	3 (4.9)	61 (100%)	3.33 (0.82)
The description of the evaluation report was appropriate from the overall standpoint.	7 (1.3)	59 (10.8)	158 (29.0)	290 (53.3)	30 (5.5)	544 (100%)	3.51 (0.81)
The rating in each evaluation item was appropriate from the overall standpoint.	8 (1.5)	83 (15.4)	155 (28.7)	258 (47.8)	36 (6.7)	540 (100%)	3.43 (0.88)

Those ratings were based on the institution's goals and objectives and then it was impossible to compare the results among the evaluated institutions. The press, however, reported the result by making the comparison among the evaluated universities as a uniformed ranking even though those universities had greatly different individualities and objectives, and a lot of misgivings to this reports were seen at the opinion survey for the self-evaluators.

On the other hand, from the questionnaire survey it was reconfirmed that the way to describe the evaluation results in the form of the "excellent point" and the "point to be required improvements" in line with the institution's goals and objectives was effective to cause the promotion of individuality and improvement. In addition to that some universities requested the evaluators to show the countermeasures to the worse points pointed out in the report in spite of the evaluation design that the NIAD-UE does not interfere the university management by taking university's autonomy into consideration and the task of the NIAD-UE is only to encourage university's self-motivated improvement. It means that the university's self-motivation for the improvement may still be low.

(2) Workload of the evaluation and self-evaluation

Even though the evaluation promoted the enhancement of university quality, both the persons engaged in making a self-report and the persons who actually evaluated the institutions expressed complaints about their great burden in the evaluation process. As for the amount of work in the university side, the vast majority of respondents answered that the evaluation imposed a heavy burden on the person in charge of the self-evaluation (Table 3). The other tasks concerning the evaluation had the same tendencies for the responses, but especially, the workload of the collection of the documents and data was answered to be large in the opinion survey.

The schedule of the evaluation influences the burden of the evaluation. Though the most of answers were "not enough" for the setting of the period of the self-evaluation, the rate of the answer was decreasing year after year because the universities had become accustomed to the evaluation work. Another factor of the heavy workload for making the self-evaluation report was that the task was concentrated at the particular persons. It seemed to be important for the targeted institutions to involve the other staff and set the evaluation work as a daily duty in order to reduce the burden on the persons in charge of the self-evaluation. By doing so it is expected that the self-evaluation process and results will be connected to the self-motivated university improvement.

As for the evaluator's work, there had been a lot of opinions on the heavy load of the document investigation and the overcrowded schedule on the first cycle. NIAD-UE was improving the evaluation process by reviewing the schedule and format. Simplification and formulation of the evaluation sheet used in the document work was contributed to the reduction of the workload and the complaints from both side decreased and the burden came to be seen as reasonable. However, reducing the workload still remains a problem for both NIAD-UE and the institutions. Especially, the data collection was heavy duty for the institutions though the evidence-based evaluation using appropriate data was the main concept of the trial evaluation. Some universities are preparing the system of the daily collection, arrangement, and accumulation of the documents and data to use for the evaluation and their own management, and then the reduction of the workload can be expected in the future.

Table 3 Workload and schedule for the self-evaluation

Item	Not so much <——————> Too much					total (%)	AV. STD
How much amount of workload did you have for making the self-report	4 (0.7)	7 (1.3)	15 (2.8)	132 (24.2)	387 (71.0)	545 (100%)	4.63 (0.67)
Item	Not agree at all <————> Strongly agree					total (%)	AV. STD
To make the self-report obstructed the daily education and research activities in your organization as a whole	50 (9.2)	113 (20.8)	196 (36.1)	144 (26.5)	40 (7.4)	543 (100%)	3.02 (1.07)
The workload for the self-evaluation was concentrated at the particular person(s)	0 (0.0)	4 (0.7)	31 (5.7)	147 (27.1)	361 (66.5)	543 (100%)	4.59 (0.63)
The period assigned to make the self-report was enough. — 1st cycle	18 (13.7)	37 (28.2)	28 (21.4)	39 (29.8)	9 (6.9)	131 (100%)	2.88 (1.18)
2nd cycle	13 (5.4)	61 (25.3)	73 (30.3)	77 (32.0)	17 (7.1)	241 (100%)	3.10 (1.03)
3rd cycle	7 (4.1)	39 (22.7)	47 (27.3)	58 (33.7)	21 (12.2)	172 (100%)	3.27 (1.07)

4.3 Outcome of the trial evaluation

(1) Enhancement of the University Quality

In the evaluation report, not only the relative level of the individual evaluation items on the basis of institutional goals, but also excellent works, elements needing improvement, and the suspected cause of problems that troubled the institution were shown. The results of questionnaire survey for the enhancement of the university quality based on the evaluation were shown in table 4.

Many self-evaluators responded that the evaluation promoted the improvement of their institution's targeted activities. They also answered it promoted the management improvement, distinctive activities, and raised faculties consciousness of education and research activities. Besides, the experience of the evaluation had possibilities to change the targeted institution's management style. It was implied by the positive results of the following questions such as "Importance of setting the goals and objectives permeated in your organization," "Importance of the self-evaluation permeated in your organization", and forth.

Most of institutions had carried out their improvements respond to these evaluation results according to the responses from the targeted institutions. Table 5 showed the number of improvement cases by degree of referring the NIAD-UE evaluation result. The cases for improvement were many and various, but the most popular improvement was establishment of systematic organization or relationship between relating organizations to introduce or strengthen the fundamental functions such as need survey, database, publicity activities, faculty development, reconstruction of curriculum and education method, development of new funding system, promotion of research projects, etc. Those improvements were considered to be an outcome of the trial evaluation.

Table 4 Enhancement of the University Quality

item		Not agree at all <————> Strongly agree					total (%)	AV. STD
This evaluation promoted (will promote) the improvement of your institution's targeted activities	UwTE	8 (1.9)	33 (7.9)	112 (26.9)	244 (58.7)	19 (4.6)	416 (100%)	3.56 0.78
	EEA	2 (3.2)	1 (1.6)	15 (23.8)	42 (66.7)	3 (4.8)	63 (100%)	3.68 (0.73)
	ERA	5 (8.2)	6 (9.8)	18 (29.5)	27 (44.3)	5 (8.2)	61 (100%)	3.34 (1.04)
Faculties' consciousness of education and research was raised by the evaluation (EEA&ERA)		7 (5.8)	7 (5.8)	50 (41.7)	53 (44.2)	3 (2.5)	120 (100%)	3.32 (0.86)
The evaluation promoted (will promote) improvement of your organization's overall management. (UwTE)		9 (2.2)	34 (8.2)	134 (32.3)	222 (53.5)	16 (3.9)	415 (100%)	3.49 (0.79)
The evaluation promoted (will promote) distinctive activities in your organization.		27 (5.0)	50 (9.3)	209 (39.1)	220 (41.1)	29 (5.4)	535 (100%)	3.33 (0.90)
More systematic evaluation was realized comparing with the former self-evaluation under the old system.		11 (2.0)	41 (7.6)	150 (27.9)	295 (54.9)	40 (7.4)	537 (100%)	3.58 (0.82)
More useful result could be earned comparing with the former self-evaluation under the old system.		16 (3.0)	44 (8.2)	197 (36.8)	252 (47.1)	26 (4.9)	535 (100%)	3.43 (0.83)
Staff in your organization fully understood the university evaluation		10 (1.9)	45 (8.4)	169 (31.5)	280 (52.2)	32 (6.0)	536 (100%)	3.52 (0.81)
Importance of setting the goals and objectives permeated in your organization.		11 (2.1)	47 (8.8)	194 (36.4)	246 (46.2)	35 (6.6)	533 (100%)	3.46 (0.82)
Importance of the self-evaluation permeated in your organization.		5 (0.9)	44 (8.2)	164 (30.7)	287 (53.7)	34 (6.4)	534 (100%)	3.56 (0.77)
Importance of systematic management on the targeted activities permeated in your organization.		8 (1.5)	36 (6.7)	179 (33.4)	271 (50.6)	42 (7.8)	536 (100%)	3.57 (0.79)
The evaluation result permeated in your organization.		14 (2.6)	89 (16.7)	272 (51.0)	146 (27.4)	12 (2.3)	533 (100%)	3.10 (0.79)
The evaluation was useful for making future plan or mid-term plan.		15 (2.8)	28 (5.3)	132 (24.8)	299 (56.1)	59 (11.1)	533 (100%)	3.67 (0.85)
Good practices could be referred in the other university's evaluation report.		19 (3.6)	100 (19.1)	261 (49.8)	134 (25.6)	10 (1.9)	524 (100%)	3.03 (0.82)

However, there were many comments from the respective institutions that not only the problems pointed out in the evaluation report, but also the process of making a self-report according to a defined evaluation framework were actually useful in finding the unknown point toward beginning improvement. It was confirmed by the interview survey. Therefore, it was made clear that both the self-evaluation process in the universities and the evaluation result shown by NIAD-UE led to an improvement of universities' educational and research activities.

From a viewpoint of management, there was a problem that the persons who did not engage in the evaluation process in the institutions had less interest in implementation

of the evaluation and the evaluation results by the results of questionnaire and interview surveys. To enhance the university quality all of the staff should act more or less toward its goals and objectives. How to involve all of the staff in the evaluation and improvement process was a problem for the next step of the effective management in universities.

Table 5 Number of the improvement cases by degree of referring the NIAD-UE evaluation result

Theme/ Category	No. of improvement case	Degree of referring the NIAD-UE evaluation result to the improvement					No answer	Total no. of cases	No. of respon-dents
		Not at all <————>Very much							
UwTE	The systems for implementation or improvement	6	13	66	235	100	7	427	425
	Methods or contents for the implementation	2	11	35	158	71	9	286	
EEA	The systems for implementation or improvement	1	1	11	44	16	2	75	65
	Methods or contents for the education or learning support	1	0	10	47	19	1	78	
ERA	The systems for implementation or improvement	1	2	5	45	21	4	78	62
	Method or functions for the research activities	0	0	8	36	11	0	55	
	Contents of the research activities	0	1	2	16	6	0	25	
Total		11	28	137	581	244	23	1024	552

(2) Promotion of the universities' accountability

To attain the universities' accountability, university stakeholders such as prospective students, employers and donors, would be enabled to utilize the evaluation results. Though the NIAD-UE disseminated the results widely to society in an easily understood form, it was pointed out that the society was not much interested in the results. The result of questionnaire survey indicated that the self-evaluators in universities thought that the evaluation did not promote understanding the universities' activities (Table 6).

It is supposed that the reason why the society did not utilize the evaluation result is that the evaluation results were not understandable for the public. Actually, the results were showed in the long sentences in the evaluation reports and they can be compared only by checking the goals and objectives that the university has held. People who are familiar with the comparable ranking and rating might not have an interest for the NIAD-UE evaluation result.

The problem embarrassing the persons concerning the universities is that the newspapers interpreted the results as a ranking of universities to make a catchy article, regardless of NIAD-UE's evaluation design that the results could not be compared among institutions. Besides, they reported only the worse points of the universities in their region even though the universities were recognized that have many good points. All of the respective persons and universities were anxious about this ill-usage, but the misunderstandings gradually disappeared as the NIAD-UE's evaluation design was gradually understood by the press.

However, the problem of the apathy of the public about the evaluation must be solved.

The NIAD-UE should improve the way to show the evaluation result. It is also needed to set up better communications between universities and society by sharing the university information described in the evaluation results.

Table 6 Degree of understanding of the universities' activities

Item	Not agree at all <————> Strongly agree					total (%)	AV. STD
People gained a better understanding of your university's activities through the trial evaluation	46 (8.6)	128 (23.9)	294 (54.9)	62 (11.6)	6 (1.1)	536 (100%)	2.73 0.82
Candidates for admission to the undergraduate and their parents gained a better understanding of your university's activities through the trial evaluation	45 (9.4)	174 (36.4)	228 (47.7)	28 (5.9)	3 (0.6)	478 (100%)	2.52 (0.77)
Candidates for admission to the graduate school and their parents gained a better understanding of your university's activities through the trial evaluation	49 (10.3)	161 (33.8)	238 (50.0)	26 (5.5)	2 (0.4)	476 (100%)	2.52 (0.77)
National and Local Government and public organization gained a better understanding of your university's activities through the trial evaluation	44 (8.3)	126 (23.9)	254 (48.2)	99 (18.8)	4 (0.8)	527 (100%)	2.80 (0.87)
Parties concerning the university (e.g. business enterprises, non-profit organization, etc) gained a better understanding of your university's activities through the trial evaluation	7 (8.7)	7 (28.5)	50 (50.8)	53 (11.7)	3 (0.4)	520 (100%)	2.67 (0.81)

5. Conclusion

From the survey results it is safe to say that one of the missions of the NIAD-UE trial evaluation, "improvement of universities' educational and research activities," was attained to a greater or less extent, but another mission, "promotion of universities' accountability," remained a distant goal. Even though the trial evaluation contributed to enhance the university quality, we should consider many problems to be solved to achieve both of the goals more effectively.

In the followings the remaining problems are described.

Problem concerning evaluation method

➢ Appropriate settings of the evaluation items that considers the characteristics of the universities and of the targeted unit for the evaluation such as institutional level, program level, department level and subject level should be considered.
➢ The evaluator should have the knowledge of the situation and characteristics of the education and research activities and management in the universities on some level.
➢ Balancing of evaluation efficiency and quality of the evaluation result is needed to be realized in pursuing the simplification and efficiency in the evaluation process.
➢ The system of the daily collection, arrangement, and accumulation of the documents and data to use for the evaluation is hoped to be established in each university.

Problem concerning the evaluators and self-evaluators
- ➢ Further improvement of the evaluator's training and development of the training for persons in charge of self-evaluation in universities are necessary.
- ➢ The communications between the targeted institutions and NIAD-UE is needed to be more closely for the effective evaluation.
- ➢ To promote the further improvement of the education and research activities in universities a lot of parties concerned should be involved to the evaluation process and promote utilization of the evaluation results.

Problem concerning publicity of the evaluation result
- ➢ The contents of the evaluation report and the way of making public are needed to be improved so that the evaluation result can be effectively used by the evaluated institutions and stakeholders.
- ➢ The NIAD-UE must increase their effort so that all of the stakeholders may have a proper understanding about the purpose of the university evaluation and the meaning of the evaluation results; especially the effort is needed to the mass communication as a medium to influence society.
- ➢ Even for the evaluated organizations, the appeal to the public is needed so that the society may have an appreciation of the organization's activities, and then they should offer more information such as the self-evaluation result, evaluation result, and improvements and quality enhancement.

To solve those problems NIAD-UE must share their experiences and lessons learned widely and construct new evaluation frameworks in cooperation with a lot of parties concerned. To realize the missions in the new university evaluation scheme, even in the process NIAD-UE needs to have collaboration with the respective institutions and the public including mass media.

Goal

Goals of the Trial evaluation

〈Quality Enhancement〉
Improvement of the educational, research, and social-service activities of universities

〈Accountability〉
Promotion of the accountability to win broad public understanding and support

Objective

Objectives that should be fulfilled to attain the goals

| Promoting university's individuality | Encouraging university's self-motivated improvement | Developing feasible evaluation | Developing fair evaluation | Making public the results in an easily understood form |

Basic framework of the trial evaluation

| Multi evaluation methods | Appropriate evaluation unit of target institution | "Goals and objectives" based | Self-report based | Peer review | Evidence based | Efficient | Transparent | Appropriate feedback and publication |

Action

NIAD-UE

Evaluation system

Organization structure for the evaluation
Committee for the university evaluation, Subcommittees, Evaluation teams, Working terms, and Clerical staff

Improvement system

Notification and dissemination of the purpose, framework and methodology of the evaluation

Evaluation process

Structure of the evaluation
Theme, Structure of the items, elements and check point, Classification of university activities, Evaluation standard and way to show the evaluation result

Implementation of the evaluation
Analysis of the self-report and documents and data to be evidence
Hearing and Site-visit

Explanation of the evaluation to the targeted institutions
Training of the evaluation to the evaluators

Making the interim evaluation report
Dealing with objections from the institutions
Publishing the evaluation report

Targeted institutions

System for self-evaluation

Implementation of Self-evaluation

External factors influenced outcomes?
Unexpected outputs or outcomes exist?

Output

| Evaluation Result and its Description in the report | Work load of evaluators in the NIAD-UE side | Work load of self-evaluators in the institution side | Accumulated experience and lessons learned Improvement of the system |

External factors

Outcom

Targeted institutions

| Grasping the activities to accomplish the goals | Description of the education and research activities | Readiness of the self-evaluation and improvement | Accumulated experiences in evaluation |
| Enhancement of education and research quality | Promotion of university's individuality | Development the consciousness of quality, continuous improvement, accountability, systematic management | Dissemination of evaluation culture |

Society

| News reporting by mass media | Understanding of the target institution's situation |
| Utilization of evaluation result by the national and local Government, by the candidates for admission, by the industrial world, etc. | Supports for the target institution | Cultivation of evaluation mind |

Attainment of the trial evaluation's goals

〈Quality Enhancement〉
Improvement of the educational, research, and social-service activities of universities

〈Accountability〉
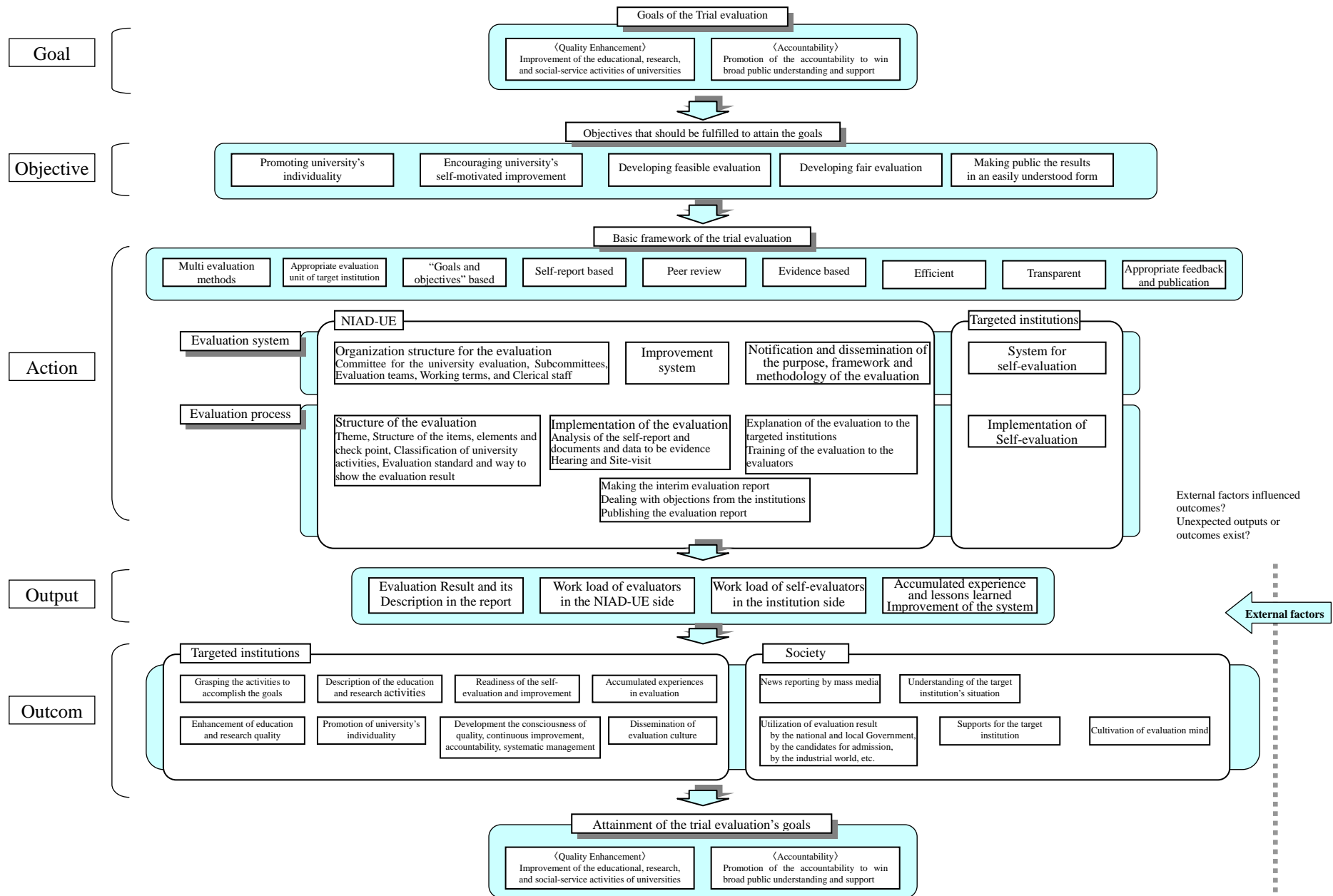Promotion of the accountability to win broad public understanding and support

Figure 2 Logic Diagram of the trial university evaluation