Quality Assurance for Higher Education in Japan

NIAD-UE
National Institution for Academic Degrees and University Evaluation

This publication is available on our website www.niad.ac.jp

Printed copies are available from: International Affairs Division
National Institution for Academic Degrees and University Evaluation

本書は独立行政法人大学評価・学位授与機構ウェブサイトで入手することができます。
冊子をご希望の方は当機構 評価事業部 国際課までお問い合わせください。

1-29-1 Gakuen-nishimachi, Kodaira, Tokyo 187-8587 Japan
Tel +81 42 307 1616
Fax +81 42 307 1559
Email kokusai@niad.ac.jp

Contents

Foreword

In today's international community, globalization is pushing countries and regions around the world into fierce competition in every aspect -- political, economic, social and cultural. As the term "knowledge-based society" suggests, in a world where people and information easily move beyond national boundaries, the competitiveness of individuals, organizations, countries and regions in the global market depends on how they can make use of a wide variety of information. In such circumstances, the importance of training people to acquire advanced knowledge and skills through education and research is increasingly recognized. Many countries are aware of this and are working to reform or reinforce their higher education to make them fit the globalized world.

Each country and university is working very hard to make its higher education go global in different ways depending on location, population, history and language. In Europe, the Bologna Process, which is aimed at creating a European higher education area, is spurring students and academic staff members to move among universities in European countries. In Asia, at the second Japan-China-Korea Trilateral Summit meeting held in Beijing on October 10th, 2009, the promotion of academic mobility in Asia accompanied by quality assurance was proposed based on recognition of the importance of mobility among universities.

International educational exchange must involve the issue of the quality of university education and research. For students studying in countries whose educational systems are different from those of the students' own countries, it is essential to receive education with more or less the same quality in any country. In other words, it is desirable that the quality of educational outcomes achieved by students meets international standards wherever the students study. Such quality assurance is expected to encourage student and other forms of academic exchange.

Since a huge amount of public money is spent on higher education, every university should be required to offer education and research whose quality is worth public spending, so evaluation and accreditation are to be implemented based on internationally agreed standards. People are now aware that proactive disclosure of information about quality assurance and fair and proper assessment of the quality of services offered by universities are very important. In addition, all countries show strong determination to enhance the competitiveness of their universities through quality assurance.

Needless to say, as universities compete globally, it is important for them to publicize themselves appropriately. Nonetheless, information given by universities alone is inadequate, so the necessity of information about university evaluation being provided by third parties is internationally recognized. The task assigned to third-party evaluation bodies is primarily quality assurance of university activities. Steady implementation of highly fair, transparent and accurate quality assurance is being called for internationally. Thus, building a third-party evaluation system that meets international standards is a responsibility of the government and a requirement for our country to earn trust from the international community.

The National Institution for Academic Degrees and University Evaluation (NIAD-UE) produced the Glossary of Quality Assurance in Japanese Higher Education to encourage a better understanding of the Japanese evaluation system, and to promote international collaborations regarding quality assurance. Moreover, in an attempt to foster, develop and solidify evaluation culture, NIAD-UE has launched a number of publications in University Evaluation Series. In 2008 NIAD-UE published the "Evaluation and Quality Assurance of Higher Education in Japan" (English version of "Development of University Evaluation Culture: Techniques of Easy-to-Understand University Evaluation", first volume of University Evaluation Series) to explain history of higher education and systems of its quality assurance.

This publication is the English translation of part of the fifth volume of University Evaluation Series. The fifth volume brings up for discussion where university evaluation and quality assurance should go in the future to learn from the drawbacks found through NIAD-UE's evaluation endeaver in the prevoius 10 years or so. Also discussed in the publication based on its experience in international partnership activities are the callenges to be tackled to help university education go global.

March 2012

Akihiko Kawaguchi
Specially Appointed Professor
NIAD-UE

Chapter 1
Accreditation, Audit, and Assessment

Third-party evaluation involves the three separate functions of accreditation, audit, and assessment. Since the emphasis on these three functions may differ with the various types of evaluation currently performed, in this chapter we will examine a couple of case studies and discuss how these functions should be incorporated to assure quality in universities or higher education.

In the evaluation of higher education, accreditation and audit both share evaluation standards and are structurally similar. In fact, some elements of auditing in Europe[1] and Australia[2] overlap with accreditation, making it difficult to distinguish the two. Therefore, the structural differences of accreditation and audit have been simplified in Fig. 1-1. In concept, the three elements of evaluation viewpoints, evaluation items, and criteria or levels form a logically connected system, flowing from top to bottom.

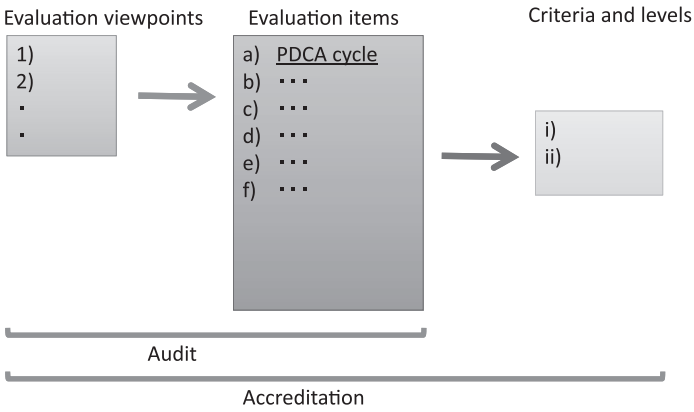

Fig. 1-1 Structures of Accreditation and Audit

In accreditation, evaluation items are established based on evaluation viewpoints, and then, minimum criteria (levels) that should be fulfilled are set for some items. This is the fundamental difference with auditing, since audits do not have any clear criteria or levels. The audit structure is designed to focus

on whether or not the PDCA (plan-do-check-act) cycle is functioning properly within the evaluation viewpoints and evaluation items. Though in reality, since some evaluation items do not always belong to the PDCA cycle, such as clarity of objectives, it is difficult to distinguish audit from accreditation.

Basically, the underlying idea of audits is that as long as the PDCA cycle is functioning properly, the evaluation items will be assured by an internal evaluation. This is based on trust between the university and the evaluation and accreditation organization that the university's self-discipline and self-assessment abilities are in full force and functioning properly. The UK has a history of universities performing quality assurance themselves, with a tradition of peer reviewers (scholars and academic staff outside the university) providing advice in the planning of standards for curriculum design or academic degrees. Therefore, an audit in the UK is based on the idea of trusting the internal quality assurance system conducted by the universities and simply checking that it is working properly.

Section 1
Accreditation

When translating the word "accreditation" into Japanese as *ninsho*, the concept may be a little vague compared to the English since the Japanese word encompasses a broader meaning. Thus, this section will explain the two meanings that *ninsho* may point to: proving qualification and certifying quality by focusing on the processes/systems of production or other activities (quality assurance). Accreditation is an action to confirm whether or not the subject (person, goods, or organization) meets the required conditions or fulfills a certain level to be qualified. *Ninsho* focusing on processes and systems is an action to confirm that production processes, as well as systems to improve these processes, are functioning properly and that they meet a certain level.

1. Accreditation to prove qualification

Accreditation to prove qualification is about proving that the subjects (organization, persons, or goods) are at a level worthy of certain qualifications. A typical example of this would be the accreditation of skills or professional qualifications. People may take exams to prove they are qualified to become,

for example, architects, accountants, tax accountants, or small and medium enterprise management consultants. Tests such as English proficiency exams and bookkeeping exams can also be categorized as this type of accreditation. These certifications can be acquired through national exams, but private organizations may also issue independent certificates.

Some goods gain accreditation through tests and verifications to prove that the quality of the product meets certain standards. A typical example of this is Woolmark, a certification based on global quality standards. There are such certifications based on international standards, and there are others that prove a certain quality level according to domestic standards. Medicines and foods also have quality assurance systems in place, but many of these take production processes into account in their verification, and thus may be better categorized as quality assurance focused on processes or systems.

There are also various types of accreditation for organizations. When starting a certain type of business, it may be necessary for the entity to apply for government registration and acquire a permit. This would show that the government has ascertained that the organization is qualified to open a business. An example of this is travel agent registration in Japan. When opening a travel agent, the person conducting the business will be screened based on standards such as whether or not the operator has a national license as a travel services manager.

2. Quality assurance focused on processes and systems

Accreditation in terms of proving qualification focuses on the state of the subject (organization, person, or goods) at the time of valuation. With goods, the focus would be on the product that results from the production process, or in other words, the output. With an organization, the focus would be on the preparation before starting operation. In either case, they focus on a certain point in time.

Quality assurance involves viewpoints and evaluation methods that are different from the above. Instead of focusing on a certain point in time, it tends to focus more on the production environment or processes to produce services or products. Therefore, rather than focusing on a single point in time, quality assurance assesses the subject in the production process over time. The most widespread quality certification in the world, one that established the basis of

this way of thinking, is the International Organization for Standardization (ISO). Here we will refer to ISO standards in order to examine concepts in quality assurance.

The first stages of quality control involved the unification of standards. ISO initially started by setting international standards for products such as screws and film speed. Then, as the service industry emerged and grew large enough to account for a substantial portion of the world trade volume, ISO decided to include the service industry in its scope. This evaluation focused on the above-mentioned output of products or services. Then ISO started to introduce standards focusing on production processes and other management systems. Examples of this are the ISO9000 quality management system standards and the ISO14000 environmental management system series.

Let us a take a closer look at the ISO14000 environmental management system standards. This evaluation is comprised of key points and requirements, and the requirements are established as questions regarding evaluation according to the key points. Furthermore, with some requirements, criteria are established as levels to be met (Tables 1-1, 1-2). This coincides with the evaluation viewpoints, evaluation items, and criteria and levels in Fig. 1-1. The point to note in the ISO14000 standards is the key point of aiming to continually achieve improvements and enhancements each year by establishing a system. The requirements are also designed in consideration of this point (Table 1-2). “Establishing a system” refers to the PDCA cycle, and continuous improvements are anticipated through the proper functioning of this cycle. If we sort the requirements according to PDCA, it may be observed that they collectively form a PDCA cycle (Fig. 1-2).

Table 1-1 Key Points of the ISO14001 Environmental Management System Standards
Source: Japan Management Association interpretation of ISO14001standards (arranged by the author)

• Importance must be placed on raising the awareness of top executives regarding environmental conservation activities. (Top executives have a key role in planning environmental policies and reviewing environmental systems.)
• The company must conduct a self-assessment of the environmental impact of its business activities.
• The company must aim for continuous improvements and enhancements each year by establishing a system.
• All details of activities must be clearly stated and documented.

Table 1-2 Requirements of the ISO14001 Environmental Management System Standards

Source: Japan Management Association interpretation of ISO14001 standards

4.1 General requirements
4.2 Environmental policy
4.3 Planning
4.3.1 Environmental aspects 4.3.2 Legal and other requirements 4.3.3 Objectives, targets, and programs
4.4 Implementation and operation
4.4.1 Resources, roles, responsibility, and authority 4.4.2 Competence, training, and awareness 4.4.3 Communications 4.4.4 Documentation 4.4.5 Control of documents 4.4.6 Operational control 4.4.7 Emergency preparedness and response
4.5 Testing
4.5.1 Monitoring and measurement 4.5.2 Evaluation of compliance 4.5.3 Non-conformance, corrective and preventive action 4.5.4 Control of records 4.5.5 Internal audit
4.6 Management review

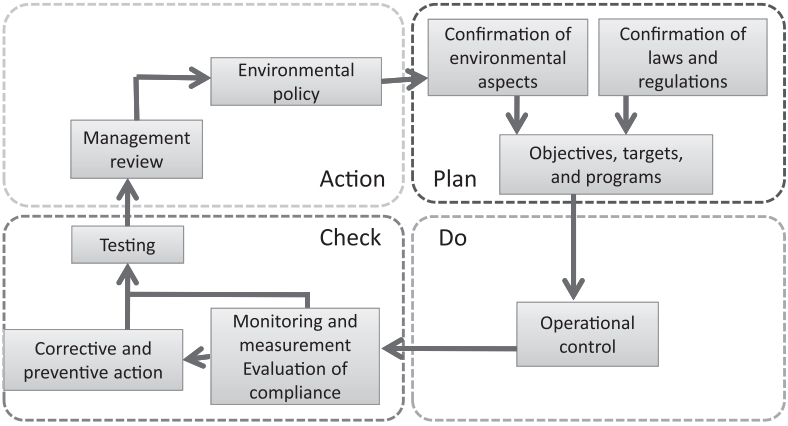

Fig. 1-2 ISO14001PDCA Cycle

Based on International Environment, Health and Safety Governance Organization material (Japan). The part above the dotted line in Plan is the process called assessment. This is ex-ante evaluation for planning, and is included in Plan in a broad sense.

Since the ISO14000 standard is a management system that takes particular interest in environmental aspects, some requirements set criteria regarding environmental impact and are checked for compliance. The method of setting criteria or levels differs according to the country, but it is said that Western countries tend to set stricter values.[3]

3. Accreditation in higher education

With the above two types of accreditation in mind, let us reassess the meaning of accreditation in higher education. Specific examples will be examined in Chapter 2 (p. 29).

Accreditation in higher education may be defined as in Column 1-1.[4,5] This definition touches on the content of educational processes, such as the acceptance of students and learning resources, and cites that accreditation in higher education values the fulfillment of these minimum standards. Therefore, the approach to accreditation in higher education is closer to quality assurance that focuses on processes and systems.

Column 1-1
Accreditation in higher education is an evaluation process to decide or reaffirm whether or not an institution or program possesses a certain level (position) or appropriateness. This is performed according to preestablished minimum standards regarding teacher qualifications, research activities, acceptance of students, learning resources, etc.

Since accreditation in higher education verifies whether or not a process is functioning properly, it is similar to auditing. In fact, some evaluation items are common with audit items that focus on systems and processes, and it is not unusual for the terms “audit” and “accreditation” to be used without clear distinction.

In quality assurance in higher education, it is important to assure the quality of academic degrees such as bachelor’s, master’s, and doctor’s, as well as professional qualifications, and a similar framework may also be envisaged for checking processes and systems to assure the quality of international validity. However, it must also be noted that levels and criteria may be influenced by regional, cultural, or political differences.

Column 1-2
Quality assurance in higher education refers to ensuring the quality of academic degrees or professional qualifications.

Section 2
Audit

The word “audit” is translated as *kansa* (audit) or *kanshi* (overseeing) in Japanese, and usually involves confirming compliance with laws, regulations, or compliance guidelines, and requesting corrections or penalties based on legal grounds if any violations are identified. Thus it suggests that a certain amount of legal force is involved. However, in the evaluation of programs and businesses, an audit refers to the act of confirming the reliability of internal evaluation or investigation. Audit in higher education also means confirming that the internal evaluation or quality assurance system is working properly. We will go onto discuss auditing in higher education after a brief overview of general audits and the meaning of auditing in program or business evaluation.

1. Audit to confirm compliance

Auditing may apply to several subjects such as accounting (financial results), businesses, or operations, but since our main concern is higher education and research projects, we will focus on audits in projects and operations. In Japan, the Administrative Evaluation Bureau of the Ministry of Internal Affairs and Communications (MIC) performs both evaluation and monitoring (audit). Thus, we will consider the differences in concepts and work methods through examples of MIC.

MIC explains administrative evaluation and monitoring as follows: MIC, as an organization specialized in evaluation and monitoring with a different standpoint from other ministries, conducts investigations focusing on compliance, appropriateness, and effectiveness regarding the work performed by ministries, in order to encourage solutions to important government administrative issues or to promote and secure the effectiveness of administrative reform according to Article 4-18 of the MIC Establishment Act; and based on the results, it encourages improvements in administrative

operations and regulations by issuing recommendations and other measures to each ministry.[6] MIC uses the word *kanshi* (overseeing), but it may be read as *kansa* (audit) in this context. What should be noted here are the points it emphasizes – compliance, appropriateness, and effectiveness. MIC basically focuses on whether or not operations are performed appropriately, in compliance with laws and regulations.

To give a specific example, in FY2008, administrative evaluation and overseeing was conducted in nursing care insurance to enhance and reinforce preventive measures against fraudulent claims, and the results were as follows. In a study examining the audit performance of 76 municipalities, it was found that quite a few municipalities had no audit at all (19 municipalities accounting for 25% of the total), and only three municipalities (accounting for 3.9%) were performing five projects to ensure the appropriate allowances necessary for care benefits. Based on these results, MIC issued a recommendation to the Ministry of Health, Labour and Welfare (MHLW) to make improvements in its system, including the way of providing know-how regarding municipality supervision of nursing care business services. Though this recommendation may not be legally binding, it is significant in that a ministry is sending a recommendation to another ministry. To receive a recommendation means that the subject has a major social responsibility to make improvements.

As such, MIC conducts administrative evaluation and overseeing by checking compliance with established laws and regulations or government policies, and sends out instructions for improvements in the form of recommendations if the performance rate is low.

Meanwhile, MIC's policy evaluation is mainly performed in two ways. The first is a meta-evaluation, which involves conducting a government-wide assessment of the policy (or measure) evaluations performed by each ministry. The second involves MIC itself conducting the policy evaluation.[7] These evaluations are conducted with attention given to necessity, effectiveness, and efficiency, and focuses on the effects resulting from the execution of the policies. Therefore, in addition to assessing the appropriateness of how the effects are analyzed, MIC identifies the necessary improvements or the elements that could be emphasized to become more effective, and then sends a recommendation to the relevant ministry.

MIC's administrative evaluation and overseeing and policy evaluation may be similar or overlap since they both focus on efficiency, but there is a fundamental difference in that administrative evaluation and overseeing places emphasis on the compliance aspect of a business, whereas policy evaluation focuses on the effects of the policy.

2. Auditing in program or business evaluation

Auditing in program evaluation or business evaluation is explained in Column 1-3. Auditing is based on the premise that the organization performs an internal evaluation, and it refers to the act of confirming the reliability of the procedures or documents used in the evaluation. In other words, auditing is not about measuring or confirming the effects of executing the program or business, but evaluating the reliability of the work performed to confirm the effect.

Column 1-3
Auditing in program is a systematic review by an independent third party of the internal evaluation of an organization or documents maintained by evaluators. The objective is to assess the reliability of the procedures and solidness of conclusions reached by the internal evaluator.

An important element in auditing is the material used in the internal evaluation. This is called the audit trail. The audit trail is not simply a stack of documents. It refers to an organized, orderly series of material related to the research and analysis work for the evaluation, including all data collected in the investigation, theoretical framework serving as the basis of the analysis, explanations of theories and models, explanations regarding data or analytical procedures, explanations regarding discovered items or conclusions, and memos written during analysis and observations. Therefore, as well as being the material used in the internal evaluation, the audit trail also serves as the material for a third-party audit.

Audit in higher education

Based on the premise that a system of quality assurance exists within universities (in other words, universities perform financial audits, operational

audits, and evaluations regarding education and research), auditing in higher education is the act of confirming the reliability of these procedures (Column 1-4). It also tries to confirm that the responsible persons are clearly defined, that the university is making improvements based on the results of the internal quality assurance work, and that this is contributing to better education and research activities. Thus, auditing in a higher education institution mainly involves confirming that the university performs internal inspections and evaluations, and that the results are leading to improvements, or in other words, that the PDCA cycle is functioning properly.

Column 1-4
Audit in higher education is an inspection of the situation or effects of the institution's internal initiatives or procedures to ensure quality (persons responsible, communication and coordination within the institution, etc.). It is often implemented on an institution level rather than program level.

The fundamental difference between accreditation and audit is that while accreditation sets minimum criteria and levels that should be fulfilled, an audit does not have clear criteria. An audit focuses on whether or not the PDCA cycle is functioning properly and does not question criteria or levels, but this is based on the idea (or trust) that a functioning PDCA cycle is capable of maintaining a certain level. However, audits in Europe[1] or Australia[2] share some evaluation items with accreditation, and so, audit and accreditation are structurally similar. As such, accreditation and audit may have overlapping areas, making them difficult to understand.

Section 3
Assessment

The word "assessment" is used in various fields such as nursing care assessment or environmental assessment, and the act it refers to may vary considerably. For example, nursing care assessment refers to collecting information to identify patients in need of nursing care or to assess potential health problems. Environmental assessment (environmental impact evaluation)

refers to studying and predicting the environmental impact of public works or other large-scale development projects and evaluating their feasibility. In other words, in nursing care, assessment refers to the act of measuring the subject to identify its state, while environmental assessment not only measures the state of the subject environment but also the effects of executing the project. Thus, assessment is a word used differently according to the field, subject, or action, and is difficult to define as a single concept.

Assessment in higher education can be defined as in Column 1-5.[7] This definition includes several acts of evaluation, the first of which is measuring the subject to identify its state and level. This alone does not include the dimension of time, and it would be an assessment of values measured at a certain time before or after the project. However, as seen in the words “input”, “process”, “output”, and “outcome” in the second half of the definition, assessment in higher education does involve the concept of time before and after the education program. Therefore, it becomes necessary to measure the state of the subject at crucial points in the process, and in this case, to explain how far the target has been achieved against the target values set at the time of planning. If it is necessary to confirm the subject’s change as a result of executing an education program or other project, the effects achieved must be confirmed not only through output but also outcomes.

Column 1-5
Assessment in higher education is the act of measuring the institution, educational program and certain structural elements. This may be qualitative and quantitative measurement according to internal and external benchmarks regarding input, process, output and outcomes, and may also accompany ratings.

Based on this definition, this section will explain the three concepts of 1) measuring the state and level of the subject, 2) conducting an evaluation to confirm the achievement level (performance measurement), and 3) conducting an evaluation to assess the effects that have developed (program evaluation).

1. Measurement of state and level

The measurement of state and level refers to assessing the state of the organization, person, or other subject. In general, this does not include the dimension of time, and refers to measuring the state of the subject at a certain point before or after the project is executed. However, as with the words “needs assessment”, the subject may be measured based on the premise that the project is to be executed. In this case, the concept of time would be included in measuring the state and level of the subject.

The organization, person, environment, or other subject may be measured for their state or level. For example, confirming whether or not the subject is at a level worthy of qualification, confirming the management situation of an organization, or confirming whether or not the atmosphere is normal. The common element in all of the above is that they do not need to focus on the process before or after the project, and do not accompany the concept of time.

However, to gauge its level the subject must be compared to some kind of guideline. This is called benchmarking, and the guidelines for benchmarking are criteria or level values. Some of the methods of setting criteria or level values are shown in Table 1-3.

Table 1-3 Methods for Setting Criteria or Level Values for Benchmarking

Best practices
The results of an organization or person with the best outcomes are set as the level value. This is suited to strong competitive environments.
General indicators
The average results of the large group to which the subject belongs to (such as the national average value) are set as the level value, and this method is the most general and easy to use. General indicators are more stable compared to level values based on best practices which rely on the results of certain organizations or persons.
Legal or certification standards
Legal standards that need to be fulfilled, such as the legal establishment standards required for schools or hospitals to open business, are set as the level value. They do not necessarily need to be stipulated by law. Certification standards required for qualifications can also be set as the level value.

When selecting a level value it is necessary to clarify what is to be confirmed. Best practice values are effective in planning strategies to survive in a competitive environment while dissociating oneself with the top executives. General indicators

are effective in assessing one's status within a large group. Laws or certification standards are effective in proving one's credibility within society.

The definition in Column 1-5 mentions that assessment in higher education "may also accompany ratings". Ratings or rankings are when a third party classifies the subject against several others. In this case, the third party assesses the state of the subject based on certain indicators and determines its rank based on those values. Indicators are selected depending on the focus in the subject, but it is generally the case that indicators for the level values in Table 1-3 are used.

The first thing to do when executing projects or programs is to identify the problems or issues of the subject person or goods (environment, animals, plants, etc.). The results are then analyzed and a project plan is drawn up and executed. In this case, even if the subject of measurement is the same, it will accompany the dimension of time, with the reference points – before execution of the project, during execution, and after execution – changing over time.

The idea of measuring the state and level of the subject within the process of higher education evaluation is shown in Fig. 1-3. The time line shows the stages from assessing the issues of the subject to the planning, execution, and completion of the project, and the actions are shown in the two rows below. The actions are divided into the following: assessing the state of the subject that directly benefits from the project, and assessing the content and management of the project plan. The latter may also include the assessment of information regarding the state of the subject.

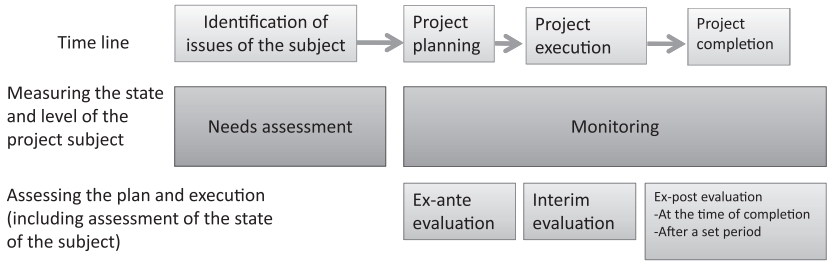

Fig. 1-3 Measuring State and Level Over Time (in the context of higher education evaluation)

In general, if the state and level of the subject are to be measured based on the premise that the project will be executed, it will involve the dimension of time. Since timing may also dictate the information required, measuring performed before the planning stage is called “needs assessment”, with a clear distinction from “monitoring” in the planning stage onwards. As input, output, and outcome are included according to the definition of assessment in higher education, the scope covers the planning stage through to completion.

Needs assessment is measuring the state and level to identify the problems and issues of the subject, as in Column 1-6.[8] The word “needs” refers to the gap between the subject’s ideal state or image and actual situation. Systematic procedures refer to the series of investigative activities to assess the state of the subject. Prioritization refers to placing the identified needs in a certain order to clarify which should be addressed first. The distribution of resources refers to the activities that should be performed and the capital and human resources that should be allocated based on this order. In other words, needs assessment provides the basic information to form plans for a project or program.

Column 1-6
Needs assessment is the series of systematic procedures for prioritization in order to develop plan or distribution of resources. They are determined by prioritized needs.

However, the volume or quality of subjects may not necessarily be the same in the needs assessment stage and the planning stage. Rather, the subjects identified in the needs assessment stage may be narrowed down in the planning stage by prioritization. This is why a line is drawn between the needs assessment stage and the planning stage onwards.

After identifying the problems and issues of the subject, solutions are considered and an action plan is drawn up. Then the plan is executed and the duration is completed, but the viewpoints or objectives of what should be confirmed will differ depending on each stage of the process. The ex-ante evaluation of the planning stage mainly involves confirming the validity of the project objective, in other words, whether or not the plan is appropriate as a method to resolve the issues, whether or not the appropriate people have been targeted, or whether or not the budget or human resources have been allocated

appropriately. The execution stage involves an interim evaluation or the action that is generally described as “monitoring”. Here it will refer to identifying whether or not the designated activities are being executed according to plan and if any are late.

The ex-post evaluation is performed at the time of project completion, but the main concern immediately after completion is whether or not the planned activities have been executed within the designated period. This is called a completion evaluation report. Sometimes the effects of executing a project or target achievement level are confirmed a short period after the project is completed. This is called program evaluation. Performance measurement and program evaluation are both ex-post evaluations, but since it is often the case that a certain amount of time is required in order to confirm the changes or impact on the subject and the effects of the project, the evaluation to confirm the state of project execution may be separated from the evaluation to assess the developing effects.

2. Performance measurement

The evaluation of the midterm objectives and midterm plans of national university corporations involves assessing how much a university has achieved in its midterm objectives and midterm plans (approved by the Ministry of Education, Culture, Sports, Science and Technology [MEXT]) within a set period. The minister is deemed as the assignor and the university as the assignee, and the relationship between the two is based on a promise in the form of the midterm objective and plan. Thus, the evaluation work confirming the level of achievement according to the contract or promise between the assignor and the assignee is described as performance measurement.

To understand performance measurement it is necessary to understand the underlying ideas of administrative reform or administrative management theories. These will be discussed in detail in Chapter 3 (p. 44), but this section will examine the basic ideas and procedures of performance measurement.

The concept of performance measurement was developed in the United States, and it is now a method employed by many OECD countries in their management and budget systems. Several definitions are used to describe its meaning, but Column 1-7 features the definition by Harry P. Hatry[9] who developed the concept. Performance measurement refers to identifying targets

in administrative execution and ensuing results, and confirming the level of achievement by measuring the indicator values. Furthermore, the efficiency and effectiveness of the relevant project are also assessed by comparing the outcome level against the execution cost, performance of sector peers, or best practices.

Column 1-7
Performance measurement is the act of periodically measuring the results or outcomes and efficiency of a service or program with applicable indicators.

In policy evaluation in the United States or international organizations, program evaluation is sometimes implemented together with performance measurement. In Japan, the distinction tends to be vague since it is often the case that *seisaku hyoka* (policy evaluation) or *gyosei hyoka* (administrative evaluation) points to both. Thus the words need to be used carefully according to different subjects or objectives; otherwise it may lead to wasted cost or inefficient work. Program evaluation identifies the social impact and effects of policies and measures through more scientific analysis, and clarifies the causal relationship between policies and their effects. Though performance measurement focuses on the increase or decrease in indicator values, it does not aim to clarify the causal connection between policies and effects; and in addition, the period is limited and the impact range is also limited. Meanwhile, program evaluation deals with policy impacts that require time in a broader scope. More on program evaluation will be discussed in the next section (p. 23).

Ten steps have been proposed as procedures for performance measurement, as shown in Table 1-4.[9]

Since the main techniques involved in the above are setting targets and measuring outcomes, we will discuss this in further detail. The outcomes of established targets can be defined as follows[9].

“Outcomes contribute to the target the program is trying to achieve and are the final benefits gained by the customer from services provided by the program.”

Table 1-4 Performance measurement procedures

Step 1: Clarification of the policy target
The targets and priorities of policies and measures should be stated clearly. However, at the target stage, the definition should be qualitative rather than quantitative to avoid being influenced by the flow of time.
Step 2: Outcome indicators and collection of indicator data
See the main text for the details of 1) definition of outcome, 2) development of performance indicators, and 3) collection of indicator data.
Step 3: Benchmarking
Benchmarks are set for comparison with collected performance indicator values. Basically, the target value based on the plan is used as the benchmark. Otherwise, generally designated standards, average values, or best practices may also be set as benchmarks.
Step 4: Performance agreement
Related parties reach an agreement regarding the target, outcome, performance indicators, data collection methods, and benchmarks. With agencies in particular, this becomes the agreed items or contract content between the governmental ministry and the agency.
Step 5: Data-based analysis
Data are collected regularly based on performance indicators and analyzed for changes or fluctuating trends in indicator values.
Step 6: Comparison with benchmarks
The performance indicator measurement results are compared against the benchmarks set in Step 3, confirming the achievement level of the objective and status among peers.
Step 7: Identification of problems, causes, and improvement measures
Problems are identified through comparisons with benchmarks. If possible, the causes are identified from the data, and points and methods for improvement are proposed.
Step 8: Submission and disclosure of the report
The performance measurement results are compiled as an annual report each year, and submitted to the ministry and disclosed to the public at the same time.
Step 9: Use of the report by the ministry
With agencies, the ministry receives the report, and from the results, considers whether or not to continue or make changes to the relevant project.
Step 10: Reflection on the budget
As with Step 9, the performance measurement results become the basis for determining the next budget plan.

What should be noted here is that the correlation of targets and outcomes is not perceived as a one-on-one relationship. Usually it is suggested that there must be several outcomes to achieve one target. The following have been pointed out as guidelines for considering outcomes.[9]

- Outcomes that a program (policy or measure) is trying to achieve (how much has been achieved by a policy addressing an issue)
- Enhancement of the quality and effectiveness of services provided by the program
- Negative secondary impact of the program
- Positive secondary impact of the program

Once an outcome is defined, indicators are developed and selected accordingly. In this case, it is crucial to develop and select figures centered on outcome indicators. Moreover, they should be developed or designed. In other words, based on the definition of "outcomes", outcomes should be estimated from the specific state of the subject (person or goods), and the most appropriate quantitative or qualitative information that embodies these outcomes should become indicators. The more the state can be described specifically, the easier it becomes to project the relevant indicators. Furthermore, one outcome is often converted into more than one indicator.

Data are collected regularly based on the developed indicators. Performance measurement is said to force many organizations to collect more data than before, and thus it is necessary to narrow down the acquirable indicators in view of the balance of cost and system. If the data cannot be found in existing government statistics or in the organization's self-accumulated data, then independent investigations will be conducted.

We have been discussing Hatry's performance measurement methodology, but this is simply a general principle that actual operations may not necessarily follow. Of the 30 OECD countries, 27% designate the outcome as the target value, but others use a combination of outcome and output.[10] Incidentally, none of the countries designate output alone. Though 46% of the countries do not have remunerative or punitive measures based on the success or failure of performance measurement results, 20% do have remunerative or punitive measures that increase or decrease budgets. Furthermore, with 16% of the

countries, performance measurement results are reflected in the payroll of the relevant organization.

Performance measurement work is in itself simple, that is, collecting and analyzing data based on indicators. However, its basic way of thinking – how it has been implemented in the appointed agency's system within the flow of administrative management reform – is important to ensure the correct understanding of university evaluation and this is discussed in Chapter 3 (p. 44).

3. Program evaluation

Program evaluation is conducted to identify the impact resulting from the execution of the project. The conventional idea in policy evaluation or administrative evaluation is that the focus should not be on input and output, but on outcome. Higher education is no exception. Even with performance measurement it is recommended to use outcome indicators to represent performance during a set period: a certain performance is identified by setting an outcome indicator and measuring changes before and after the execution of the project. However, strictly speaking, simply measuring changes in the indicator values before and after the project may not exactly represent the impact of the project itself. Elements other than the execution of the project may have had an influence in raising the indicator values, as with the following example of students.

How is it possible to confirm that lectures and training programs have yielded the desired effects on students? The most typical way would be to conduct exams to confirm how much the students have understood. However, unless it is possible to determine whether student understanding was a result of attending lectures, or was inherent to the students regardless of the lectures, or was acquired through other educational opportunities, it may not be declared as the direct result of the relevant lectures.

Then how should the impact of a project be measured? Let us examine the meaning of impact and then discuss the method of measuring impacts.

<u>What does "impact" mean?</u>

If the impact of a project cannot be described simply by changes in the indicators before and after the project, then how should it be gauged? The answer is that it is necessary to compare the changes with and without the

project. Fig. 1-4 is a graph showing changes in the indicators when the project is executed and when it is not. The change when the project is executed is difference A. Even if the project is not executed, the subject changes with time and that change is difference B. Therefore, the impact of executing the project would be acquired by subtracting B from A, resulting in difference C. Since C is the direct impact of executing the project, this is called “net effects”.

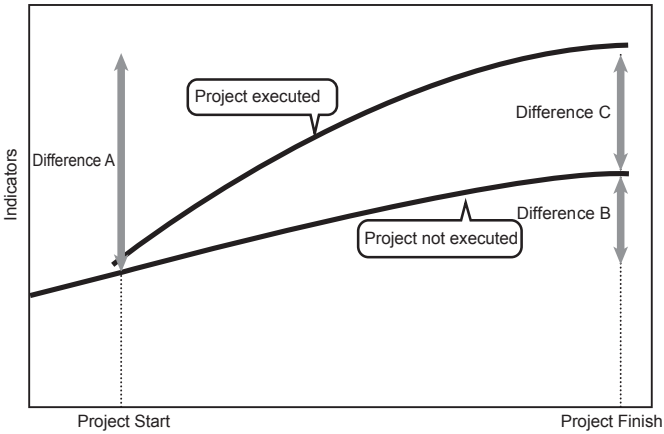

Fig. 1-4 Net effects

Net effects are measured not only by changes in the subject that was provided with services from the project, but also by a comparison with the subject without services. Thus, the subject group with services from the project is called the treatment group, while the other without services (no project) is referred to as the control group. To make it possible to compare the treatment group against the control group, efforts must be made to maintain both groups in the same quality, size, and environment.

The subject that executed the project (treatment group) is measured for a certain period for indicators to show an impact, and the subject without the project (control group) is also measured for the same period for the same indicators. If an impact is recognized from executing the project, there should be a difference between the two groups. The difference would emerge as a discrepancy in indicator values. However, this difference is still not the “net effects” of the project, because it is possible that the evaluation itself may be affecting the treatment group subjects. For example, a subject person who is

conscious of being studied by an evaluator may try harder than usual. Therefore, in theory, the true net effects would be subtracting the indicator value of the subject without the project from the indicator value of the subject with the project, and then subtracting the indicator value converted from the impact of evaluation.

Moreover, the subject is affected not only by being evaluated but also by the external environment besides the project. Therefore, the effects of the external environment must be subtracted when discussing the impact of the project. However, this point may be resolved to a certain extent by creating a comparative subject group. In other words, since the treatment group and the control group not only have the same quality and amount, but are also placed under the same environment, and the same conditions would create a similar external environment, the difference between the two groups would cancel out the impact of the external environment.

However, though this may be possible theoretically, it is not easy to create this kind of condition in reality. Unlike guinea pigs in a laboratory, humans live in various economic, social, and cultural environments, and thus it is impossible to create environments of identical nature. This is the most objective and scientific way to assess the impact of a project, but it is also necessary to take into account the limitations of studying humans who live in diverse environments. It may also be difficult to deliberately create a group with no project or services from moral and humanitarian considerations. In this case, an alternative method to a comparative subject group must be found.

Method to measure impact

Various methods have been devised to measure impact or net effects. The method of creating two comparative subject groups, staying true to the idea of net effects, is called the experimental approach. Meanwhile, the method that does not create a control group is called the quasi-experimental approach. Here we will examine typical methods of each: the random experiment model and the regression discontinuity model.

The random experiment model involves choosing persons from the subject population through random sampling. Here “random” means that if the same opportunities and services are provided, it is highly likely that the subjects will react in the same way in a probabilistic sense. The subjects selected by random

sampling are then randomly divided into two groups. When doing so, two groups of the same size are created, with members of the same nature and characteristics, and subjected to the same conditions as much as possible.

The net effects of executing the project, or in other words, changes that would occur to the subjects if services were provided, is defined specifically and translated into indicators that would express them in the best manner. The indicators for the two groups are confirmed, and the two groups should display the same value at this stage.

The project is then executed and services are provided to one group (treatment group). The other group has no services since the project is deliberately not executed (control group). The indicators for each group are measured at this stage. The treatment group with services will be expected to show higher values than the group without, and the difference is measured. This difference may be said to be the impact of the services, or the impact of the project.

The quasi-experimental approach measures the net effects through an alternative method without creating a control group, and the method said to be the most reliable is the regression discontinuity model.

The basic procedures for this model are as follows:

- A regression line is drawn for the subject group.
- Based on this regression line, the group is divided into two at a certain threshold.
- A service is provided to one of the two groups.
- A regression line is drawn again a certain period later.
- If there is a discontinuation in the regression line, the distance of the discontinuity may be assessed as the net effects of the service.

An example is shown in Fig. 1-5. A student scores 2.5 or above in the first-term performance and is awarded the principal's prize. A regression line is then drawn for the second-term performance. The figure shows a discontinuity in the regression line of the awarded student scoring 2.5 or above and the regression line of the unawarded student with a lower score. Since the scores are 2.4 and 2.6, the difference or distance of discontinuation is 0.2. This difference is perceived as the impact of winning the principal's prize.

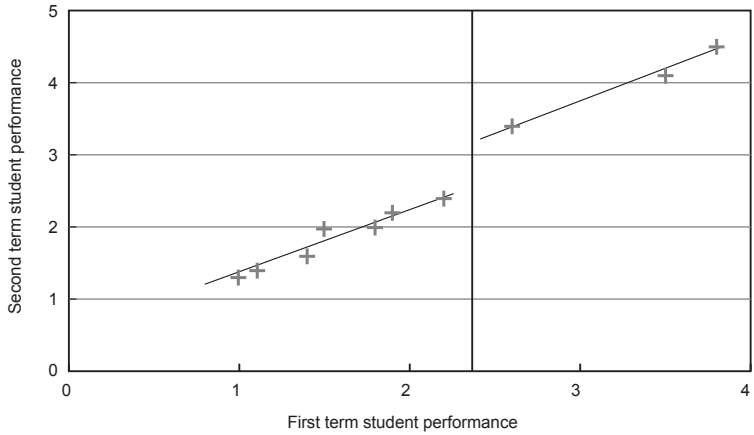

Fig. 1-5 Impact of School Awards on Student Performance

These are the basic ideas regarding methods to measure the impact of a project as scientifically and objectively as possible. Here we have tried to gauge the impact of the project itself as much as possible, and to explain the causal correlation between the project and the outcome through the term "net effects". Obviously, since our real lives or educational environment are affected by various external elements besides the project, the impact cannot be described as completely pure. However, for significant projects that involve large-scale budgets and human resource investments, using these methods to try and identify the impacts contributes to the objective of achieving accountability. It also serves to provide important information for future project development or non-development.

Notes:

1. Amourgis, S., et. al. (2008). *Programme-oriented and institutional-oriented approaches to quality assurance: new developments and mixed approaches*, European Association for Quality Assurance in Higher Education (ENQA).
2. University evaluation conducted by the Australia University Quality Agency (AUQA) is called "audit", but it is performed based on a model focusing on the process from plan execution to improvements called ADRI (approach-deployment-results-improvement), and this has many points in common with process evaluation in quality assurance (accreditation).

3. Since some of the criteria in the ISO14000 requirements in Japan are set relatively loose, a separate environmental report may be prepared because ISO14000 alone would not be sufficient in proving the environmental performance of the organization. According to ISO cases, there have been frequent occurrences of misconduct or trouble among ISO14000 certified companies in Japan that have focused on system checks, highlighting the difference with U.S. companies that set criteria.
4. Rout, M. (2008). "Degree standards 'must be ranked'", *The Australian Higher Education Supplement*, February 20.
5. Woodhouse, D. (2004). *The quality of quality assurance agencies, Quality in Higher Education 10 (2)*, Routledge, pp. 77-87.
6. Ministry of Internal Affairs and Communications (MIC), Administrative Evaluation Bureau (2005). "What is administrative evaluation and monitoring?", MIC website (in Japanese)
   (http://www.soumu.go.jp/main_sosiki/hyouka/hyouka_kansi_n/index.html)
7. This is also divided into two types. The first is an evaluation to secure unity in policies that are common in several ministries and perceived as being in need of evaluation from the government-wide view of securing unity. The second is an evaluation to secure coherence or comprehensiveness in policies that are related to several ministries and perceived as being in need of evaluation from the point of view of comprehensive promotion.
8. Witkin, R.B., Altschuld, W., J. (1995). *Planning and Conducting Needs Assessment A Practical Guide*, SAGE.
9. Hatry, P.H. (1999). *Performance Measurement Getting Results*, The Urban Institute Press.
10. Organization for Economic Co-operation and Development (2005). *Modernising Government: The Way Forward*, OECD Publishing.

Chapter 2
Certified Evaluation and Accreditation

All graduate schools, universities, junior colleges, and colleges of technology, whether national or private, are obligated to undergo third-party evaluations periodically under the School Education Law. This mandatory review scheme is Certified Evaluation and Accreditation (referred as CEA hereafter). CEA is conducted by an evaluation organization certified by the Ministry of Education, Culture, Sports, Science and Technology (MEXT) (hereinafter "CEA organization") regarding the comprehensive state of their education, research, organizational operations, and facilities so that the results may help enhance the level of their education and research.

The three objectives of CEA are 1) assure quality and accreditation, 2) encourage improvements and enhancements, and 3) achieve third-party accountability regarding activities of education and research in graduate schools, universities, junior colleges, and colleges of technology. Therefore, CEA should encompass the functions of accreditation and audit mentioned in Chapter 1. There are two different categories of CEA: one is institutional CEA which evaluates the state of the institution as a whole; and another is CEA for professional graduate schools. Universities, junior colleges, and colleges of technology are required to undergo institutional CEA by a CEA organization at least every seven years, and professional graduate schools are required to undergo evaluation at least every five years.

An evaluation organization must be certified by the Minister of MEXT. This certification is given to an organization which meets certain requisites including its standards, method and framework for assessing fair and accurate evaluation. In this way, the standards, method and framework of evaluation differ from each organization. As of October 2011, the National Institution for Academic Degrees and University Evaluation (NIAD-UE), Japan University Accreditation Association (JUAA) and Japan Institution for Higher Education Evaluation (JIHEE) had been certified as institutional CEA organizations for universities. CEA is to be carried out, at the request of an institution, in accordance with the standards for CEA set out by an implementing organization.

The following is a description of institutional CEA for universities (Section 1) and CEA for law schools (graduate schools) implemented by NIAD-UE.

Section 1
Institutional certified evaluation and accreditation

The institutional CEA performed by NIAD-UE as of FY2010 consists of 125 universities, 13 junior colleges, and 60 colleges of technology. Since NIAD-UE's standards for institutional CEA are largely the same for universities, junior colleges, and colleges of technology, we will be discussing universities from here on.

The standards of CEA for universities are comprised of eleven criteria (FY2005-2011, Table 2-1), which apply to all universities, and optional evaluation items. NIAD-UE's CEA assesses the overall state of a university's activities centered on the main educational activities in a formal university program. However, research activities are important along with educational activities for universities; furthermore, universities also perform activities to give back knowledge to society through both education and research by partnering or interacting with local communities or industries. Thus, taking into account the objectives of helping to improve the university's various activities and achieving accountability, optional evaluation items have been established besides the eleven standards. These assess the state of research activities or conditions of educational services offered to those other than full-time students, which are difficult aspects to assess simply through educational activities. These optional evaluation items assess related activities according to the university's wishes.

Table 2-1 Standards of institutional certified evaluation and accreditation for universities

Comparison of first cycle and second cycle.

First Cycle: FY2005-2011	Second Cycle: FY2012-2018
1: Purpose of the University	1: Mission of the University
2: Education and Research Structure	2: Teaching and Research Structure
3: Academic Staff and Education Supporting Staff	3: Academic Staff and Teaching Supporting Staff
4: Student Admission	4: Student Admission
5: Academic Programs	5: Academic Programs
6: Effectives of Institutional Performance	6: Learning Outcomes
7: Student Support	7: Facilities and Student Support
8: Facilities	8: Internal Quality Assurance System of Teaching and Learning
9: Internal Quality Assurance System	9: Financial Base and Management
10:Finance	10:Public Information on Teaching and Learning
11:Management	

Each of the stipulated eleven standards is divided into a number of contents, and basic viewpoints are set accordingly. Whether or not the standards are met is judged not by individual viewpoints or content, but in a comprehensive manner by each of the eleven standards. When all the standards (apart from the optional evaluation items) are fulfilled, the university is deemed as fulfilling the CEA standards as an institution. If the university fails to meet even one standard, it would be seen as failing the CEA standards for university as a whole. However, no university has been judged as failing these standards as yet.

NIAD-UE believes that the objectives of CEA are not achieved by simply judging whether or not the standards are met. If among the fulfilled standards a certain initiative is recognized as exceptional or recognized to meet a certain standard, but is in need of improvement, these points will be noted in the evaluation and accreditation report. Noting these points of excellence or areas of improvement is important for the second objective of helping improvements and enhancements.

The optional evaluation items, unlike the eleven standards, are not interested in judging whether or not criteria are met; they are more concerned with evaluating the achievement level of the objectives established by each university. This is evaluated in the four levels of “excellent”, “good”, “satisfactory”, and “unsatisfactory”. Furthermore, the reasoning supporting the assessment, points of excellence, or areas of improvement are also noted.

First cycle of the institutional CEA ends 2011 and second cycle starts from 2012 to 2018. To start the second cycle, NIAD-UE modifies standards of institutional CEA for universities (Table 2-1). Major change in standards towards the second cycle is to introduce the concept of learning outcomes. It is also important to evaluate universities internal quality assurance system of teaching and learning and disclosure of the information on teaching and learning. In addition to these changes, results overview in English will be made public and “International Activities for the Teaching and Learning” will be added to the optional item of evaluation.

1. Have the objectives of certified accreditation been achieved?

NIAD-UE’s motto is to provide open and evolutionary evaluation. Therefore, we have conducted studies each year to verify how the three objectives of certified evaluation and accreditation have been achieved.

NIAD-UE conducted signed, multiple-choice (five levels), written questionnaires at institutions (universities and junior colleges) subjected to institutional CEA from 2005 to 2008 (Table 2-2) and external evaluators (Table 2-3). This questionnaire covered a wide scope, from the content of evaluation and accreditation to methods and outcomes after the evaluation and accreditation. Here we will focus our discussion on points such as accountability and internationalization in certified evaluation and accreditation.

Table 2-2 Questionnaire results of institutions subjected to institutional certified evaluation and accreditation (excerpt)

Answers collected: 68 out of 70 evaluated institutions (97%). Excerpt from "Verification report regarding certified evaluation and accreditation" (http://www.niad.ac.jp/n_hyouka/jouhou/index.html). Figures represent the percentage of the total number of evaluated institutions that responded to the survey. 5: Strongly agree ～ 3: Neutral ～ 1: Strongly disagree

About NIAD-UE's certified evaluation and accreditation report regarding activities of teaching and research, etc.

Question	5	4	3	2	1
Adequate for assuring quality	19	74	7	0	0
Useful for making improvements	22	68	10	0	0
Helps gain and encourage understanding and support from society	16	57	24	3	0
Gives new perspectives	9	44	47	0	0

What kind of impact or effects do you think NIAD-UE's certified evaluation and accreditation results would have on activities of teaching and research

Question	5	4	3	2	1
Help gain an overall assessment	25	65	10	0	0
Help identify future issues	19	74	7	0	0
Raise awareness in the education and research activities of academic staff	7	51	40	1	0
Encourage improvements in the institution's overall management	7	66	25	1	0
Encourage improvements	13	69	16	1	0
Assure quality	15	59	26	0	0
Help gain understanding and support from students	4	29	63	3	0
Help gain widespread understanding and support from society	4	44	47	4	0

Table 2-3 Questionnaire results of certified evaluation and accreditation evaluators (excerpt)

Answers collected: 182 out of 241 evaluators (76%). Excerpt from "Verification report regarding certified evaluation and accreditation" (http://www.niad.ac.jp/n_hyouka/jouhou/index.html). Figures represent the percentage of the total number of evaluators that responded to the survey. 5: Strongly agree ～ 3: Netrual ～ 1: Strongly disagree

About NIAD-UE's certified evaluation and accreditation overall regarding activities of teaching and research of universities (junior colleges)

Question	5	4	3	2	1
Assures quality	16	64	18	2	0
Encourages improvements	16	66	16	2	0
Helps gain and encourage understanding and support from society	8	51	36	4	0

Evaluation about the self-assessment report submitted by the university (junior college)

Question	5	4	3	2	1
Easy to understand	6	55	29	9	1
Standards for evaluation and accreditation and other content were described in an appropriate manner	4	60	30	5	1
Necessary material used as premises were quoted or attached	6	53	31	10	1
Provision of reference information would have been helpful for document analysis	5	23	43	26	3

From the submitted self-assessment reports it is observed that universities and evaluators have different degrees of understanding regarding such aspects of appropriateness and explicitness of content. Quite a few of the universities stated that they found it difficult to collect and select material as attached documents for the self-assessment reports. On the other hand, evaluators pointed out inadequacies and insufficiencies in the self-assessment reports and requested improvements in presentation. While it is true that these issues are gradually being resolved as universities accumulate experience in evaluation, it is undeniable that the degree of understanding regarding the clarity of self-assessment reports or the appropriateness of back-up material has become more varied among universities compared to when evaluations first started. An analysis of the effort put into evaluation work suggests that the daily accumulation of material and data required for evaluation will be important in

future. This problem is not limited to evaluation work; major improvements are also necessary in terms of communicating information to society.

The figures suggest that institutional CEA has produced significant results for assuring quality and helping improvements. In contrast, the objective of achieving accountability to society still remains an issue. Unfortunately, it is hard to say that sufficient results have been seen in the efforts to influence or impact the understanding and support of existing students, potential students, or society. Thus it is necessary to continue considering measures, including disclosure methods, to encourage people to understand and support the content of evaluation reports. For example, though evaluation and accreditation reports or self-assessment reports have been disclosed to the public, less than 70% of the evaluated universities felt the media coverage was appropriate, and hence there is the need for efforts to gain more understanding.

2. Issues in certified evaluation and accreditation

Two points must be added to the issues stated above. The first is the necessity for universities to accurately assess their resources. Obviously, an institution should be able to assess its own resources, but CEA has revealed that in some cases performances may not be sufficient. CEA is conducted based on NIAD-UE's standards for evaluation and accreditation, but it also takes into account the objectives and targets set by each university. This is a device to encourage uniqueness in each university through CEA. Universities must set objectives and targets based on self-assessments of their resources. If objectives and goals are set without sufficient assessment, they may end up being very vague or simply general content. As a result, the objectives and targets may not be able to convey the uniqueness or character of the university. Thus, besides any obvious lack of base material, this may be why evaluators find self-assessment reports difficult to understand.

The second problem is the insufficiency in achieving accountability. Universities are accountable to stakeholders, and it is necessary to recognize that there is an extremely diverse range of stakeholders in university's education and research – students, their families, future employees, academic staff, university managers, and others. Policy planners are also stakeholders under the present climate where higher education policies are deemed important. Naturally, the quality recognized by each stakeholder is different. When discussing the quality

of a university, for example, students will think of the university's facilities or how beneficial education and research will be for future job opportunities. The students' families will hope for academic achievement or job opportunities for their child. Employers will focus on the abilities and competence of the graduates (or students who have completed the course). Academic staff will direct their attention to the classes and learning processes. University managers will focus on the outcomes as an institution. Furthermore, policy planners will look at the effectiveness of the policies. Since each stakeholder defines quality from a different standpoint, it is impossible to discuss quality through a single concept, and thus important to communicate information with each stakeholder in mind. This is an issue that involves both the evaluation organizations, which communicate evaluation results, and the universities, which communicate information regarding their activities of education and research.

3. Sending certified evaluation and accreditation results abroad

In recent years, it is often heard that when universities make an agreement with other universities abroad, they are asked to send third-party evaluation results. In a knowledge-based society, it is essential to send information not only from the university itself, but also from a third-party evaluation organization associated with the university's quality assurance. In such a global trend, a major mission of NIAD-UE regarding its evaluation business is gaining international confidence in quality assurance. This may not be achieved by simply producing English translations of evaluation results. Since higher education systems differ in each country, it is essential to have a good understanding of these differences before being able to send accurate quality assurance information; otherwise it would be meaningless.

NIAD-UE produced the Glossary of Quality Assurance in Japanese Higher Education to encourage a better understanding of the Japanese evaluation system when communicating quality assurance in higher education, and to promote international collaborations regarding quality assurance. This glossary was compiled as part of ongoing cooperation with the UK Quality Assurance Association (QAA) and is a list of definitions of terminologies used in the Japanese higher education system, quality assurance system, and NIAD-UE evaluations in English and Japanese. Furthermore, NIAD-UE published the "Evaluation and Quality Assurance of Higher Education in Japan" (English

version of “Development of University Evaluation Culture: Techniques of Easy-to-Understand University Evaluation” University Evaluation Series) to explain systems that include evaluation culture, evaluation concepts, certified evaluation and accreditation, and national university corporation evaluation. As such, we believe that the foundation has already been laid for communicating information regarding quality assurance in higher education to counterparties abroad, and that we are now at the stage where information regarding quality assurance of education conducted by each organization should be dispatched proactively.

The requirement internationally is basically an assurance of academic degrees or professional qualifications, in other words, information regarding accreditation. If a student were to graduate (or complete a course) from a university faculty (or academic unit of a graduate school), it must be possible to recognize the student’s academic achievement, skills, and abilities. An international student would need information on the learning outcomes that may be anticipated by attending a certain university. Of course, this information must be communicated from the university itself, but an evaluation institution must also assure the quality of content in terms of the learning outcomes gained (or may be gained) by attending a university.

To address these social needs, it is necessary to improve the quality assurance system for the next evaluation cycle of institutional CEA. Thus, we suggest setting a distinction between the functions of audit and accreditation in institutional certified evaluation and accreditation to a certain extent.

An audit is an evaluation of the university as a whole to confirm whether or not internal quality assurance systems or improvement systems regarding teaching and learning are functioning properly. The next five may be considered as evaluation items.

1. Mission, vision, and objectives of the university regarding the quality of education
2. Efforts of the university in trying to realize its mission, vision, and objectives
3. Method of assessing the achievement level of the mission, vision, and objectives
4. Efforts being made for improvements and enhancements

5. Execution and responsibility for assuring internal quality and making improvements and enhancements (including suitability of established standards)

The second function, accreditation, refers to assuring the suitability or quality regarding established standards or objectives/targets set by the university. The size and organizational structure of universities vary considerably, and to perform an accreditation it is necessary to analyze the state of education and research in the university faculties and graduate school units. This is an evaluation with a particular focus on teaching and learning outcomes, and the following three may be considered as evaluation items. If the basic data regarding these items are publicly disclosed, for example, through a database, this would achieve accountability to society.

1. Execution systems, contents, methods, or other aspects of teaching (including suitability of established standards and suitability of academic staff for courses)
2. Academic achievements (including student evaluation)
3. Job opportunities or careers after graduation or completing courses (including evaluation by related parties)

Based on the above consideration, NIAD-UE has just started the second cycle of institutional CEA with new standards shown in Table 2-1 (p. 30).

Section 2
Certified Evaluation and Accreditation for Law Schools

Professional graduate schools offer practical education to train highly specialized professionals. These schools were established from the recognition and awareness that conventional graduate schools were not necessarily responding to the diverse and growing needs for training professionals – training personnel equipped with special knowledge and abilities capable of international performance, reeducating working professionals to upgrade their specialized abilities, or addressing social emphasis on qualifications. They had been more focused on training researchers.

In general, these institutions, rather than being called "professional graduate schools", are often called graduate schools of law, accounting, business (MBA or MOT), or education – names that give a more specific description of their content. Professional graduate schools are obligated to undergo certified accreditation by a CEA organization at least every five years. This section examines the issues and future outlook from the first round of CEA for law schools that is almost complete.

1. Certified accreditation

Law schools are categorized as a type of professional graduate school, but they are unique in that they receive accreditation when recognized as meeting accreditation standards in third-party evaluation. Thus it may be said that certified accreditation is all the more important with law schools. The evaluation criteria of CEA for law schools are stipulated in an extremely detailed manner by related laws and regulations.[1] The evaluation method also must be of a standard capable of performing an accurate certification stipulated by the Law on Coordination of Graduate Law School Education and the National Bar Examination. Furthermore, professionals with practical legal experience must participate in the evaluation work. The CEA results are notified by the Minister of Education, Culture, Sports, Science and Technology (MEXT) to the Minister of Justice. The evaluation organizations currently conducting CEA for law schools are the Japan Law Foundation (JLF), Japan University Accreditation Association (JUAA), and NIAD-UE. Already 68 out of 74 law schools (23 national universities, 2 public universities, and 49 private universities) have undergone CEA as of March 2009. (Table 2-4.)

Table 2-4 Number of evaluated law schools by the accreditation institutions

	Japan Law Foundation	NIAD-UE	Japan University Accreditation Association	Total
FY2006	2 (0)	-	-	2 (0)
FY2007	11 (1)	9 (4*)	2 (0)	22 (5)
FY2008	14 (6)	16 (2)	14 (9)	44 (17)
Total	27 (7)	25 (6)	16 (9)	68 (22)

As of March 31, 2009. Figures in brackets represent the number of universities found inadequate.
*Three universities were certified as adequate in follow-up evaluations in FY2008.

What should be noted here is that nearly a third of the universities were assessed as inadequate. The main reasons for this assessment in the NIAD-UE's accreditation are as follows:

1. The credits that can be registered for courses exceed the designated maximum.
2. The number of students in some basic law courses (attending at the same time) largely exceeds the average.
3. There are problems in terms of objective and strict performance evaluation and credit certification.
4. In the selection of new students, the method of identifying students who are yet to study law and those who have already studied law is inappropriate.
5. Some courses are not instructed appropriately because of a mismatch of academic staff with unsuitable achievements of education and research.

Thus, reasons for failing adequacy may range from the selection of new students to the organization of academic staff. The content is also spread over a wide scope, from issues that may violate laws and regulations to others that fail to meet the independent levels required by NIAD-UE.

2. Verification results

After completing an evaluation, NIAD-UE sends out questionnaires to the evaluated law schools to analyze the state of execution. An overview of the results is as follows.

The structure and content of the standards and interpretation guidelines have been assessed as appropriate by the law schools and evaluators in terms of assuring quality, encouraging improvements, and gaining the understanding and support of society (students, families of students, companies, and other related parties) regarding educational activities. The fact that the standards are centered on educational activities has also been assessed as appropriate. However, a certain number of both the law schools and the evaluators said some standards overlapped in content, and more than 50% said some items were difficult to evaluate. Thus it is necessary to clarify the standards and interpretation guidelines through briefings and workshops, and also to review

the appropriateness of the descriptions in the standards and interpretation guidelines with the next accreditation cycle in mind.

Regarding the appropriateness and explicitness of the self-assessment reports submitted by the law schools, it has been observed that there are differences in understanding among law schools and evaluators. Quite a few of the law schools experienced difficulties in collecting and selecting material as attachments for the self-assessment report, and many evaluators pointed out inadequacies, insufficiencies, or areas of improvements in presentation. While it is anticipated that these issues will be gradually resolved as law schools accumulate experience in evaluation, it is necessary to provide better explanations regarding the required materials and data for evaluation.

Both document analyses and site visits have been viewed positively by the law schools and have also been met with approval from evaluators. Many have commented that site visits in particular have been extremely effective in gaining a common understanding of the state of educational activities.

Regarding the accreditation reports, law schools have said that they are appropriate overall, being sufficient in assuring the quality of educational activities while also appropriate in terms of attainment of goals and objectives and actual situation of each law school, and that the explanations have been basically easy to understand. Evaluators have also shown approval that the document analyses and site visits have been reflected in the accreditation results, and that the structure and presentation of results in accreditation reports are also appropriate. However, it is hard to say that the effects and impact on the understanding and support from students (including future students) and society have been sufficient, and efforts are required to increase societal awareness of CEA for law schools performed by NIAD-UE.

For overall opinions on accreditation, many of the law schools commented that NIAD-UE's accreditation process was precise and careful, based on clear standards, and just as anticipated. Many also said they would like to see a publication compiling case studies, featuring excellent or unique initiatives. This suggests that many of the law schools felt that this accreditation process had been useful in assessing the state and issues of their educational activities, and had produced effects or impacts such as helping improve educational activities or assuring quality in educational activities.

3. Improvement measures to enhance the quality of law school education

The fact that quite a few law schools failed in their adequacy, combined with the underperformance of law school graduates in the new bar examination, has prompted comments that the quality of law school graduates is insufficient, and questions have been raised regarding education in law schools. The Central Council for Education Subdivision on Universities Law School Special Council (hereinafter "Special Council") has performed investigations to examine the actual state of each law school, reviewed the opinions of related organizations, and performed interviews with legal professionals to reach the following view of the current situation.

For the new law school system as a whole, many law schools are making solid progress in planning teaching programs to bridge the gap between theory and practice in order to fulfill their anticipated role in judicial reform. Regarding the overall talent and abilities of law school graduates, they have been regarded by persons involved in instructing legal apprentices as good as or better than those before. On the other hand, however, taking the certified evaluation and accreditation results or examination results for judicial apprentices into account, the following problems have been identified in law school education or with some of the law school graduates. It has been pointed out that these need to be resolved swiftly. [1]

1. Some graduates have been recognized as lacking sufficient knowledge in the fundamental understanding or legal thinking of basic law.
2. Some graduates have been recognized as lacking sufficient abilities of logical expression.
3. The content of basic education of legal practice in law schools is inconsistent.

The Special Council has summarized wide-ranging improvement measures for quality assurance in law school education such as 1) securing the quality and diversity of new students, 2) assuring the quality of graduates, 3) providing better education systems, and 4) establishing evaluation systems focused on quality.[2] For CEA, the Special Council requests improvements in the evaluation standards and methods to place more emphasis on quality. It also suggests that the evaluation standards establish the following as important evaluation items

to assure the quality of law school education: implementation status of unified minimum standards for testing aptitude, status of strict performance evaluation and graduate certification (including the achievement of common targets), performance and competence of academic staff in teaching and research, and career paths of graduates (including bar examination results). Furthermore, for the judgment of "inadequacy", the Special Council requests that the evaluation standards and methods be reviewed according to these important evaluation items to ensure that law school operations do not cause misunderstandings or confusion for new students who wish to enroll in law schools or for society in general. NIAD-UE is reviewing the evaluation standards or methods for the next accreditation cycle according to the Special Council report.

4. Future outlook

The number of law school graduates who passed the new bar exam fell below expectations, and while major law schools in Tokyo are competing for the top spot, smaller regional schools are struggling. Already there is talk of elimination and consolidation of law schools. Meanwhile, there are still extremely serious issues of regional differences, such as uneven distribution with only one lawyer per population of more than 30,000, or depopulation with extremely few legal firms in branch areas of regional courts. There may also be many problems that cannot be resolved by law schools alone. However, there is a vision that made law schools necessary, and law school staff and related persons must emphasize this point. It is essential for each law school to set clear objectives and targets based on a thorough assessment of their location requirements and human/material resources, and to present the outcomes of their activities according to their objectives and targets to society.

Today, it is said we are in the "age of outcome evaluation", or in an age where much emphasis is placed on outcomes. Thus it is necessary to have a thorough understanding of the word "outcomes". The results produced by a law school (within the organization) through input (financial, human, and material resources invested to execute educational activities) and action (acts or work of investing input to produce output: drawing up curriculums, holding courses, conducting end-of-term examinations, etc.) are called "output". Outcomes refer to the results including effects and impacts of the law school's various activities on the subjects (outside the organization), such as the actual achievements of

the students, or the final skills and knowledge acquired by students. Therefore, though the success rate of the bar examination is of course one of the outcomes, there should also be other outcomes anticipated by each law school. These outcomes need to be clearly communicated to society, and this may help law schools gain social awareness regarding the uniqueness or presence of each institution.

Notes:

1. Central Council for Education Subdivision on Universities Law School Special Council "Improvement measures to enhance the quality of law school education (report)" (April 17, 2009), The Ministry of Education, Culture, Sports, Science, and Technology website (in Japanese) (http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo4/houkoku/1261059.htm)

Chapter 3
National University Corporation Evaluation

The evaluation of national university corporations' achievements against their midterm objectives and midterm plans is a classic example of applying the idea of performance measurement to university evaluation. As explained in Section 3 of Chapter 1 (p. 19), performance measurement is a way of evaluation used to verify to what extent an agency (incorporated administrative agencies, for example, in Japan) has achieved the objectives that it agreed to attain in a certain period of time. Understanding why national university corporation evaluation falls into the category of performance measurement requires understanding how national universities became corporate entities and how the newly-launched national university corporation system works.

In this chapter, we will take an overall look at the theory of New Public Management (NPM) and one of its embodiments – agency systems – to add an international perspective. Then we will describe the background to the incorporation of national universities, the characteristics of the incorporation process, and the similarities and differences between the national university corporation system and the incorporated administrative agency system. We will also explain the basic design of the national university corporation system and the role played by evaluation in the system.

Section 1
New Public Management

New Public Management (referred to as NPM hereafter) is a theory of administrative reform and administrative management developed and adopted in countries such as Britain and New Zealand since the mid-1980s. NPM aimed at downsizing the public sector, improving the quality of public services, and achieving efficient management in administrative organizations.

In Japan too, administrative institutions had studied NPM since the late 1990s. When inaugurated in 2001, Prime Minister Junichiro Koizumi's administration embarked on structural reform, adopting in earnest administrative reform policies based on NPM; also introduced were the Private Finance Initiative (referred to as PFI hereafter) and competitive sourcing to make use of

market mechanisms. Behind the introduction of these systems was, as in other countries, the necessity of scaling down the executive branch of the government and operating administrative services more efficiently. In Japan in particular, the government thought that fiscal spending needed to be reduced through drastic administrative reform in order to overcome disadvantages, such as an aging society with a low birth rate and huge fiscal deficits, and create a sustainable society.

1. Basic concept of NPM

The basic concept of NPM is to introduce systems that can tap market mechanisms, such as privatization, outsourcing of governmental services to the private entities, PFI, and competitive sourcing, in order to encourage private-sector entities to contract to provide public services. In administrative organizations, NPM is aimed at urging administrators to perform in economical, efficient and effective ways, evaluating their performance from the perspective of public satisfaction, disclosing the results to the public, and listening to public voices so that such opinions will be reflected in the processes of policy-making in public services. The ultimate goals of NPM are to downsize the executive branch, deliver public services more efficiently, and improve their quality. The characteristics of NPM can be summarized as follows[1-3]:

- Controlling the executive branch by evaluating its performance and outcomes in exchange for giving it wider discretion over the use of business resources
- Making use of market mechanisms as much as possible
- Controlling the executive branch from the customers' point of view (treating the public as the customers of public services)
- Transforming the executive branch into a more easily controllable organization or executive agency (simplifying its hierarchy)

Comparing NPM with conventional administrative management or bureaucracy makes its characteristics clearer. As for management methods, traditional bureaucracy is based on laws and rules, whereas NPM puts importance on performance and outcomes. As for organizational structures,

the former is a vertical and hierarchical command system, whereas the latter adopts a system where the hierarchy is simplified, and discretion is given to the executive agency, which performs based on agreements with the mid-term objectives and the mid-term plan.

Therefore, in these two systems, jobs are carried out differently. Conventional bureaucracy adopts a vertical system where each division specializes in one particular job, but NPM does not adhere to the vertical system as much as bureaucracy does because it places importance on improving the quality of services and getting jobs done more efficiently. Also, the former uses private businesses only to a limited extent, thinking that public services will not go with market mechanisms, whereas NPM tends to take advantage of market mechanisms more actively, which results in more outsourcing.

What was behind the introduction of NPM, a theory of administrative management? Firstly, there was inefficiency involved in an overgrown bureaucracy. In general, as an organization becomes larger, extra burdens – whether official ones or not – will increase within the organization. The way of administrative management that goes through all the formalities, including rules and procedures, also imposes more burdens and causes inefficiency. Another reason is said to lie in the asymmetric possession of information within and between the administrative organizations that contain a planning division and an executive division, and since on-the-spot information tends to gather at the executive division, this division will accumulate more knowledge and expertise. Thus, when decisions are made to implement policies, preference will be given to the executive division's values and ways of thinking as a result, rather than intentionally. In the type of work where its purpose and performance are difficult to quantify, such a tendency is said to be greater. That can also happen between the cabinet, which decides on government policies, and each ministry, which implements them. People sometimes criticize the bureaucracy for its intervention in decision-making processes and its manner of putting top priority on its own interests, and that can partly be attributed to the asymmetric possession of information. All those problems would ultimately be against the interests of the public. Sovereign power is supposed to reside with the people, and the government is supposed to make sure of this. If each ministry or each

division of a ministry acts in its own interests, however, people's needs would come second, which would not only cause a mismatch between their real needs and public services provided, but also increase costs with delivery of excessive services.[4]

2. Various systems based on NPM

A number of systems have been implemented based on the NPM theory to reduce the amount of administrative services through outsourcing or to make executive branch as independent executive agency to run administrative organizations more efficiently. In the administrative organizations, as stated above, their hierarchical structures have been slimmed down and discretion has been given to the executive agency; at the same time, an evaluation system that values outcomes has been introduced, reinforced and thoroughly implemented to make its responsibility clear.

Other measures have been taken to decouple some of the administrative services from the executive branch by adopting market mechanisms. A classic example is the privatization of public services through deregulation and regulatory reforms. Even public services and functions considered best provided by the public sector have begun to be contracted out to private sector firms and nonprofit organizations through PFIs and the school voucher system. In Japan, too, PFI has been introduced as a means to tap the private sector; also implemented are Japan's own measures, such as the designated administrator system. In addition, some countries have introduced executive agency systems where public services and functions considered to be non-marketable are assigned to independent organizations under the condition that their discretion and responsibility are clearly stated. Japan's incorporated administrative agency system is the equivalent of that.

Competitive sourcing is implemented in the United States, Britain and Japan; it is aimed at allowing both the public sector and the private sector to bid on an equal footing for certain administrative functions or services and assigning them to those who make proposals designed to provide efficient and quality services. In some cases, before an executive agency system is introduced, competitive sourcing of the targeted administrative functions or services will be carried out.

Section 2
Executive Agency System

The concept of the incorporated administrative agency system influenced the government's designing of the national university corporation system. The basic framework of the concept was provided by the executive agency system, which was first introduced by Britain in 1988 when Margaret Thatcher was prime minister. In this section, we will explain how executive agencies are defined, what areas they cover, the basic ideas about how they are run, and the key system factors of agreements and evaluation.

1. What is the executive agency system?

The executive agency system is a system or framework aimed at improving operational efficiency in administrative functions and the quality of administrative services by setting up independent institutions to deliver administrative services. In Britain, employment, dismissal, appointment, salaries, and working hours of the administrative management staff, and also how to use information devices, used to be decided by the central government in accordance with rules under the centralized political and administrative system, which meant that those administrators had little discretion. This system was said to be causing inefficiency in operating administrative services, thus the British government separated the function of providing public services from each department of the government and assigned it to newly-established independent agencies in order to give more discretion to administrators and improve administrative efficiency. At the same time, the duty and responsibility of the agencies needed to be clear. Executive agencies were to conclude agreements with their parent department and treasury on their budget and objectives they had to achieve in a certain period of time; they also had an obligation to explain to what extent they had fulfilled those objectives at the end of the period. Japan adopted the concept of this British executive agency system to create an incorporated administrative agency system.

What public services do executive agencies offer? Public services are defined as something that the public sector needs to continue to provide because it is difficult for the private sector to do given the nature of the services and little attention placed on efficiency because of non-competitive environment and

etc.; but agencies independent from the government should provide them from the standpoint of efficiency. In other words, they are something that straddles both the public sector and the private sector, so obviously they are not easy to define.

Let us take a look at how executive agencies operate in terms of both budget and job. Budget ceilings tend to be strictly capped, but the budget is not tightly restricted in how it is spent, unlike the budgets of central government departments. Thus, each executive agency is given a certain degree of discretion as to how to spend its budget in spite of the budget ceiling. In Britain, executive agencies are allowed to draw up a budget covering more than one fiscal year.

The concept of contract is applied to the business of executive agencies: they are to clearly define what outcomes they will have to produce in a certain period of time and then sign an agreement on that with their parent department. In other words, government departments are the assignors and executive agencies are the assignees. Thus, the former needs to check whether the latter is doing what they agreed to do, or fulfilling their duties. This is done by evaluation and assessment. What they have achieved is not measured by the amount of resources, such as budgets, human resources and systems, put in to conduct their business or by how the programs are being implemented; it is measured by what outcomes have been produced. In addition, since people are the customers of administrative services, the outcomes should also be evaluated from the perspectives of what impact the programs have on people's lives and to what extent they are satisfied with them.

2. Role of evaluation in the executive agency system

When you look at the executive agency system you will find that evaluation plays an important role. Government departments and executive agencies are related not only as assignors and assignees, but also as planners and plan implementers. Thus it is highly likely that the executive agencies will accumulate more on-the-spot information and expertise than their parent government departments do, causing an asymmetric possession of information between them. With this and the belief that public services are uncertain by their very nature because they do not go well with market mechanisms, executive agencies are thought, in theory, to be easily tempted to act in their own interests. In other words, the executive agency system itself contains the same problem

as the bureaucracy, caused by the asymmetric possession of information. As a measure to prevent such a problem, performance measurement is adopted to evaluate executive agencies. For them to improve the efficiency and quality of administrative services requires not only monitoring their activities by measuring their performance, but also giving incentives or imposing penalties on them based on the results of the measurement. The British executive agency system stipulates in the agreements that if the results of performance measurement show that the executive agencies have achieved more than their objectives, pay raises and bonuses shall be offered; if they fail to meet their targets, however, it is possible that their budgets will be significantly slashed or they will be reorganized or abolished.

Section 3
Basic Design of the National University Corporation System

Japan's incorporated administrative agency system is in line with Britain's executive agency system introduced under the Thatcher administration, and the incorporation of national universities has a strong connection with national university reform and administrative reform. In this section we will explain the background to the policy on the incorporation of national universities, the similarities and differences between the national university corporation system and the incorporated administrative agency system, the basic design of the national university corporation system, and the role played by evaluation in the system.

1. Background to the policy on the incorporation of national universities

The incorporation of national universities was triggered by the basic plan for downsizing national administrative organizations and improving their efficiency, which was decided by the cabinet in April, 1999.[5] The plan was aimed at reforming all administrative institutions and decreasing the number of civil servants; it also included the transition of subsidiary bodies and public corporations under each ministry to incorporated administrative agencies. Turning national universities into incorporated administrative agencies is mentioned in the plan as follows:

"(1) The matter of the incorporation of national universities will be examined as part of university reform and settled by 2003 with the autonomy of universities taken into consideration." (Section 2, Chapter 2: Matters Related to the Incorporation of Administrative Agencies, in the basic plan for downsizing national administrative organizations and improving their efficiency)

As of April, 1999, the incorporation of national universities was suspended for further study, while the incorporation of other institutions had been decided. The cabinet seems to have concluded that universities should not be treated the same way as other institutions and due consideration was needed before they were incorporated.

In response to the above-mentioned administrative reform plan, the Minister of Education, Culture, Sports, Science and Technology announced the "Structural Reform Policies for National Universities" in June, 2001. This policy contained three major points: 1) drastically implement the reorganization and consolidation of national universities, 2) implant market-oriented management practices into national universities and transform them into incorporated entities, and 3) introduce elements of competition guaranteed by third-party evaluation into the management of universities and help the TOP 30 selected national, public, and private universities to be among the world's best universities. The policy clearly stated that national universities would be incorporated and that an evaluation system would be introduced. At that point, however, a third-party evaluation was intended to be carried out to determine the amount of funds to be allocated to each university through the Center of Excellence (COE) and other research subsidy programs.

In June, 2001, the "Structural Reform of the Japanese Economy: Basic Policies for Macroeconomic Management" was decided by the cabinet.[6] The Basic Policies stated that national universities were to enhance their global competitiveness by increasing their independence through the incorporation process and introducing market-oriented management practices.

In March, 2002, the final report concerning the image of national university corporations was released by a research council of the Ministry of Education, Culture, Sports, Science and Technology (MEXT).[7] This was the blueprint for the national university corporation system. The report emphasized that the incorporation of national universities must be beyond the standpoint of

administrative reform. This can be demonstrated by the following sentence of the report:

"This matter should be examined beyond the perspective of administrative reform, which involves, for example, the outsourcing of administrative functions and the improvement of operational efficiency; it should be examined from the perspective of helping universities become more vigorous and internationally competitive by pushing ahead with the ongoing university reform, which focuses on promoting the advancement of education and research, encouraging universities to develop distinctive characters, and revitalizing the administration of universities."

The report is in line with the "Structural Reform Policies for National Universities" and the Basic Policies, both announced in June, 2001, in terms of recognizing the need to enhance the competitiveness of national universities by turning them into corporate entities and introducing an evaluation system. The "Basic Policies for Economic and Fiscal Policy Management and Structural Reform 2002," decided by the cabinet in June, 2002, stated that MEXT would start incorporating national universities and converting their academic, administrative, and other staff to non-civil servant status preferably in fiscal 2004.[8] The conversion was approved by the cabinet. In November, 2002, the "Guidelines for Formulation of the FY 2003 Budget" was decided by the cabinet,[9] which stated that structural reform of universities would be driven by the incorporation of national universities and other measures to put them in a competitive environment that would encourage them to be among the world's best universities; then in July, 2003, the National University Corporation Law and other related laws were enacted.

We have taken a general look at the government's policies related to the incorporation of national universities. This move was triggered by the large scale of administrative reform, including reducing the number of civil servants and incorporating administrative agencies. However, a line was drawn between the incorporation of national universities and that of other administrative agencies because due consideration was given to the autonomy of universities, or to the point of the matter that academic freedom, guaranteed in Article 23 of the Constitution, should include the autonomy of universities too. Therefore, though

it is true that discussions about the incorporation of national universities were generated by the administrative reform policy, it would be more appropriate to understand that four years had been spent in deliberating and examining it as part of university reform in order to respect the autonomy, independence, and self-reliance of universities.

2. Characteristics of national university reform

In this subsection we will analyze the characteristics of national university reform based on which national universities were incorporated by referring to the report concerning the image of national university corporations, compiled by a research council of MEXT.[7]

The first characteristic is that each university was incorporated independently under the same legal system. As stated in the report, even though the idea of incorporation was within the framework of the incorporated administrative agency system, how it should be done had been examined with each university's individual character taken into consideration. As a corporation each university has been thrown into a competitive environment where they must offer their own appealing teaching curriculum and research programs.

The second is the clarification of vision. Each university is required to make clear their philosophies and direction of reform.

The third is a substantial expansion in the discretion of universities. Discretion has been expanded through the conversion of university staff to non-civil servant status and large-scale deregulation of the country's various systems. Since the national university corporation system is within the framework of the incorporated administrative agency system, the executive discretion of universities has been expanded as long as it is exercised to achieve the midterm objectives agreed between universities and the minister of education.

The fourth is the establishment of a management structure. Incorporated national universities have wider discretion and more independence and self-reliance; at the same time, they assume greater responsibility. Thus, the report states that who has authority over and responsibility for the administration of universities should be made clear. This provides the basis for establishing a decision-making body centered on the president and an in-house audit system that includes outside experts.

The fifth characteristic is the participation in university management by

outside experts. They take part directly in the management of universities as directors or members of the administrative council.

The sixth is the implementation of evaluation and much wider disclosure of information. The report frequently mentions the implementation of evaluation. In line with what the "Structural Reform Policies for National Universities" had pointed out, the report stated that a competitive environment would be created by introducing an evaluation system and disclosing information more vigorously in order to encourage universities to improve the quality of their education and research. The report also stated that evaluation results would be reflected in the calculation of management expenses and other grants.

3. Basic framework of the national university corporation system

The National University Corporation Law and other related laws, which came into effect in October, 2003, are products of the discussions on university reform. Fig. 3-1 describes the basic framework of the national university corporation system.[10] It shows the organizational structure of a national university corporation inside the box and the major stakeholders involved in the operation of the national university system outside the box. The president, board of directors, administrative council, education and research council, and auditor, shown inside the box, constitute the decision-making body of a national university corporation. Under these lie the faculties, graduate schools, administration bureaus, and other organizations, though these are not shown in the figure. As major stakeholders, the Minister of Education, Culture, Sports, Science and Technology, the National University Corporation Evaluation Committee, and the National Institution for Academic Degrees and University Evaluation (NIAD-UE) are shown. National university corporations have many other stakeholders, but the three parties are shown in the figure because they are thought to have a direct influence on the operation of the national university corporation system.

The Minister of Education, Culture, Sports, Science and Technology (MEXT) shall appoint the presidents nominated by national university corporations (para. 1, Article 12 of the National University Corporation Law); the minister shall also appoint auditors (para. 8, Article 12). National university corporations are required to set the midterm objectives to be achieved in a period of six years and to make them public. The draft is prepared by universities,

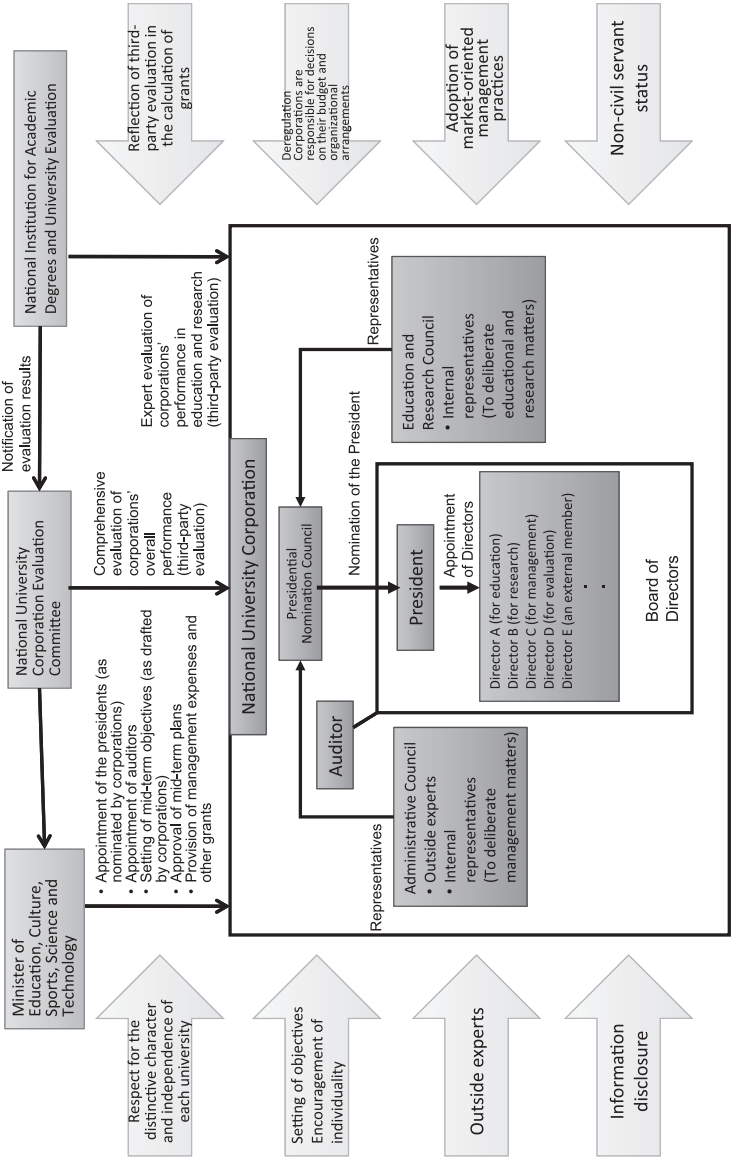

Fig. 3-1 Basic framework of the national university corporation system under the national university law

Drawn up based on the outline of the national university corporation system compiled by the Ministry of Education, Culture, Sports, Science and Technology (MEXT)

but the law stipulates that the minister shall set the midterm objectives to be achieved by national university corporations (Article 30). National university corporations are also required to compile a midterm plan for meeting their midterm objectives and obtain the approval of the minster (Article 31), and the evaluation of their performance based on the midterm objectives and midterm plans shall be carried out by the National University Corporation Evaluation Committee (Article 9).

Now you can see that the minister of education and national university corporations are related as appointer and appointees, or as assignor and assignees. Although national universities are now corporate entities, they are only granted executive discretion as to how to attain the midterm objectives set by the minister; ultimate responsibility for the attainment of those objectives rests with the minister of education. This is in line with the incorporated administrative agency system.

4. Differences between the national university corporation system and the incorporated administrative agency system

The incorporation of national universities is compared with that of administrative agencies because, as stated above, the relation between the minister of education and national university corporations is similar to that between ministers in charge and incorporated administrative agencies, and a portion of the stipulations of the Act of General Rules for Incorporated Administrative Agency are applied to the National University Corporation Law (Article 35 of the Law). Nonetheless, as indicated by the independent existence of the National University Corporation Law, the national university corporation system and the incorporated administrative agency system are two different systems. So what is the difference between the two?

The first difference is respect for education and research and respect for the autonomy of universities. In the incorporated administrative agency system too, the government is to take into consideration the administrative independence of agencies. In the national university corporation system, the characteristics of education and research, namely independence, self-reliance, specialization, and time-consuming processes are to be respected as well.

The second concerns executive bodies. How to organize, run, and manage incorporated administrative agencies is basically left to their own discretion. In

the case of national university corporations, by contrast, the major constituents of a corporation's executive body, namely its board of directors, administrative council, and education and research council, and their basic roles, are stipulated in the law (Fig. 3-1).

The third difference is about midterm objectives. In the case of incorporated administrative agencies, the ministers in charge set the midterm objectives for them and direct each agency to pursue these objectives. For national university corporations, however, the minister of education is to hear and take into account what they have to say before setting the midterm objectives for them. This is clearly demonstrated by the way the present system is run; midterm objectives and midterm plans are formulated by national university corporations themselves and are then given the green light by the minister of education.

The fourth is the appointment and dismissal of the head of a corporation. The heads of incorporated administrative agencies are to be appointed by the ministers in charge. In the national university corporation system too, the president of a university is appointed by the minister of education; however, the appointment is made based on the nomination by the presidential nomination council of the university. In addition, while the ministers in charge can dismiss the heads of incorporated administrative agencies for deterioration in their performance or other reasons, the minister of education can dismiss the president of a national university corporation only at the request of the presidential nomination council. These can be seen as other examples of respect for the autonomy of universities.

And the fifth is about evaluation. In the incorporated administrative agency system, the commissions on evaluation of incorporated administrative agencies set up under each ministry are to carry out the evaluation of agencies and notify the results to the Commission on Policy Evaluation and Evaluation of Incorporated Administrative Agencies (CPIAA) of the Ministry of Internal Affairs and Communications (MIC). In the case of the national university corporation system, however, the National University Corporation Evaluation Committee of MEXT is to conduct the comprehensive evaluation of national university corporations, respecting the results of NIAD-UE's evaluation of their education and research activities. They considered that the specialized knowledge and skills were required to evaluate education and research. This is the reason why National University Corporation Evaluation Committee

of MEXT requested NIAD-UE to conduct those evaluation. MEXT is to notify CPIAA of the evaluation results and make the results public. Another characteristic of the national university corporation system is that expert evaluation of the education and research activities of national university corporations is to be carried out not every year, but every six years.

Section 4
Evaluation of Midterm Objectives/Plans and Performance Measurement

We explained at the top of this chapter that the evaluation of national university corporations' achievements against their midterm objectives falls into the category of performance measurement. What is the reason for this? We will seek the answer, providing an explanation of how corporation systems work.

Performance measurement is aimed at seeing achievements by defining the objectives and outcomes of administrative programs and measuring them by indices. The point is under what system the performance measurement is carried out. As stated above, performance measurement is used in Britain's executive agency system and Japan's incorporated administrative agency system, and it may involve incentives and penalty – for example, the budgets of agencies could increase or decrease depending on the evaluation results. Even though some differences between the national university corporation system and the incorporated administrative agency system exist, both national university corporations and incorporated administrative agencies are to pursue the objectives set by the ministers in charge. These are the objectives of business administration that national university corporations are to achieve in a period of six years, and to what extent they have been met is to be verified by evaluation. The evaluation results are to be reflected in the calculation of grant for management expenses. Given the mechanism of the system, it can be said that the evaluation of national university corporations' achievements against their midterm objectives is a performance assessment.

Nonetheless, a difference in performance measurement implemented in Japanese national university corporations and in U.S. and British governmental institutions needs to be pointed out; as stated above, objectives are, in practice, set by the national university corporations themselves, and although MEXT

presents several objectives concerning education and research that should be part of the midterm objectives (Table 3-1), no details are given. Thus it would be appropriate to say that MEXT only provides an outline.

Table 3-1 Objectives concerning improvements in the quality of education and research of the midterm objectives of national university corporations

Major objectives	Middle objectives	Minor objectives
Objectives concerning education	① Objectives concerning educational outcomes ② Objectives concerning what to teach ③ Objectives concerning systems to provide education ④ Objectives concerning assistance for students	Specific objectives under each of the middle objectives
Objectives concerning research	① Objectives concerning levels and outcomes of research ② Objectives concerning arrangements for systems to conduct research	Specific objectives under each of the middle objectives
Objectives concerning partnership with society and international exchange	① Objectives concerning partnership with society and international exchange	Specific objectives under each of the middle objectives

Source: Guidelines for the Performance Report, compiled by NIAD-UE

Performance measurement is defined as setting indices to indicate outcomes and measuring them regularly.[11] According to this definition, objectives need to be explained using such outcome indices. In reality, however, many national university corporations use qualitative descriptions rather than numerical targets to explain their midterm objectives. As well, many objectives contain outputs, implementation processes, or inputs rather than outcomes. The fact that some areas of research and education are not suited to quantitative explanation is the reason behind this, and also that some of the outcomes and effects hardly materialize over a period of six years.

Nonetheless, as seen in the discussions about learning outcomes, attention is being paid to measuring learning outcomes in quantitative ways by policy-makers in education at home and abroad. It would be worthwhile for educational institutions to reconsider what outcomes they could produce based on the objectives they set and whether there are any outcomes that could be explained using indices. This would make the objectives clearer and also increase the accuracy of evaluation.

Benchmarks are used in performance measurement, and to see to what extent the objectives have been reached, benchmarks, or numerical targets set beforehand, are used as a yardstick. In some cases, though, in order to see the levels of their own performance and outcomes, these are compared with best practices, mean values, or standards stated in the laws. Nevertheless, few universities specify their objectives or show numerical targets. Since setting specific numerical targets requires a fair amount of work to collect information and data, more opinions as to how objectives should be described need to be heard from, among others, those concerned in universities.

1. Overall picture of evaluation in the midterm objective period

National university corporations and inter-university research institute corporations (unless otherwise stated, referred to as national university and inter-university corporations hereafter) are to have their performance during a six-year period to attain their midterm objectives evaluated by the National University Corporation Evaluation Committee of MEXT (referred to as the Corporation Evaluation Committee hereafter). As for the evaluation of education and research activities, the Corporation Evaluation Committee is to request the National Institution for Academic Degrees and University Evaluation (NIAD-UE) to conduct the evaluation and respect the evaluation results. At the request of the Corporation Evaluation Committee, NIAD-UE is to carry out the evaluation of the corporations' education and research activities (referred to as education and research evaluation hereafter) from an expert point of view, present the results to the Corporation Evaluation Committee, and make them public. The Corporation Evaluation Committee is to conduct a comprehensive evaluation of the corporations' overall performance during the midterm objective period, respecting the results of NIAD-UE's education and research evaluations (Fig. 3-2). Incidentally, a single-year evaluation is conducted by the Corporation Evaluation Committee, not NIAD-UE.

Since education and research are the core activities of national university and inter-university corporations, NIAD-UE's education and research evaluation plays an important role in the system for evaluating national university and inter-university corporations. Education and research evaluation, carried out with the administrative independence and self-reliance of national university and inter-university corporations and the characteristics of their education and

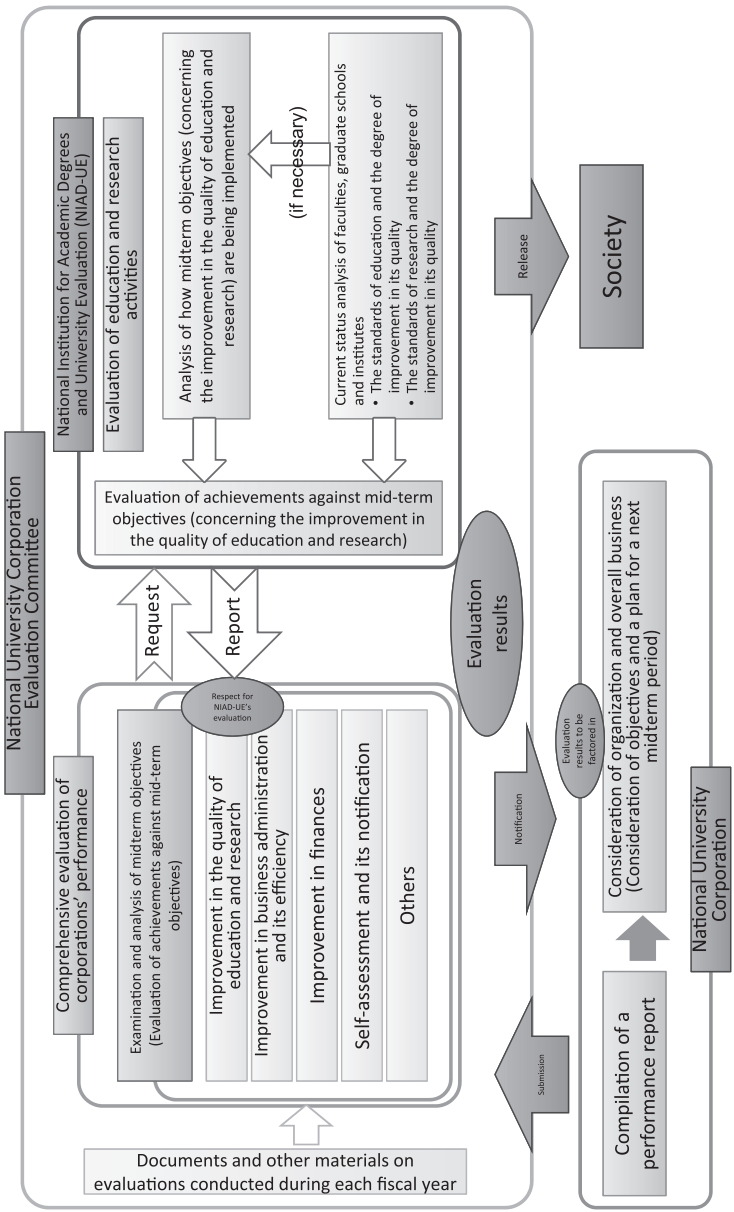

Fig. 3-2 Whole picture of the evaluation of national university corporations' performance during a midterm objective period

research taken into consideration, are aimed at maintaining and improving the standards of their education and research and also helping them develop in diverse and individual ways. At the same time, through the whole process of its education and research evaluation, NIAD-UE is expected to fulfill its responsibility to explain clearly to society how national university and inter-university corporations are performing their education and research activities.

The evaluation of national university and inter-university corporations has the following three purposes: 1) to have the evaluation results serve as a guide to formulating objectives and plans for the next midterm period; 2) to have the evaluation results reflected in the calculation of grant for management expenses; and 3) to fulfill the responsibility to provide an explanation to society. In terms of 1) and 2), in particular, the Corporation Evaluation Committee decided that the evaluation of national university and inter-university corporations' performance during the first midterm objective period would be conducted in the fifth year after the incorporation process started. So NIAD-UE carried out education and research evaluations in fiscal 2008. In addition, it will finalize the evaluation results after the midterm objective period expires.

The basic policy and implementation structure and methods of the education and research evaluation are explained in detail in the Guidelines for Performance Report, the Evaluation Guidelines, and the Evaluation Manual.[12] The Guidelines for Performance Report provides information needed for national university and inter-university corporations to compile a performance report. The Evaluation Guidelines and the Evaluation Manual are handbooks for NIAD-UE's external evaluators, but are available on its website to ensure the transparency of evaluation. As well, as for questions from corporations and its answers to them, NIAD-UE puts together Q&A lists when necessary and makes them public on its website to allow all corporations share the information.

At the request of the Corporation Evaluation Committee (Table 3-2), the performance report submitted by each university, used for NIAD-UE's education and research evaluation, is composed of a midterm objective achievement report, which covers the entire corporation, and current status reports on faculties, graduate schools, and institutes, which look at each organization of the corporation (Fig. 3-3).

Table 3-2 Requirements from the National University Corporation Evaluation Committee to NIAD-UE's education and research evaluation

1. Education and research evaluation is to be conducted with a view to
   i) being referred to in evaluating national university corporations' achievements against their midterm objectives;
   ii) having the evaluation results serve as a guide to formulating objectives and plans for the next midterm period;
   iii) evaluating national university corporation's attainment of their midterm objectives concerning education and research, referring to NIAD-UE's analysis of the current status of education and research organizations (referred to as faculties, graduate schools, and institutes hereafter) of each national university corporation.
2. Education and research evaluation is to be implemented in the fifth year (in fiscal 2008) with a view to having the evaluation results reflected in the formulation of objectives and plans for the next midterm period and in the calculation of management expenses and other grants.

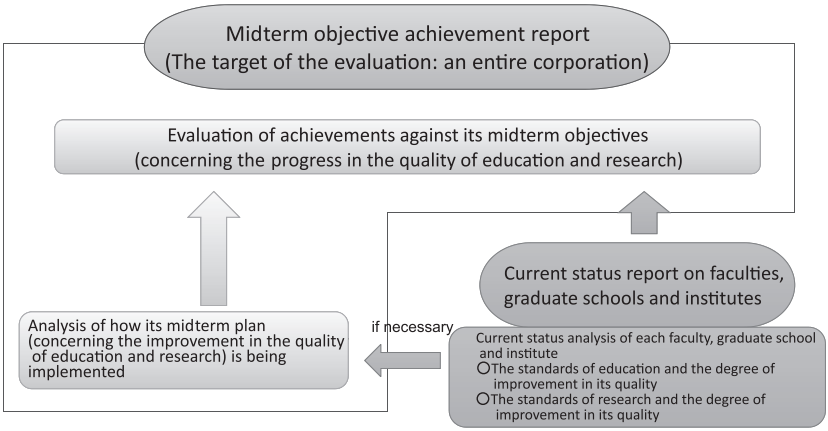

Fig. 3-3 Elements of a performance report used for education and research evaluation

Education and research evaluations were conducted by document analyses and site visits. The document analyses were carried out based on the current status reports on faculties, graduate schools, and institutes, and the midterm objective achievement report submitted by each university. After compiling the results of the document analyses, site visits were made to follow up on the document analyses. During the site visits, NIAD-UE interviewed those concerned in the universities under evaluation, checked submitted documents and data, and also interviewed the students, graduates, and teachers to grasp how the universities' education and research activities were being implemented.

The contents of the evaluation report that NIAD-UE presented to the Corporation Evaluation Committee and made public are as follows. The evaluation reports on each national university corporation and inter-university research institute corporation are available on NIAD-UE's website.[13]

1. Results of the evaluation of achievements against midterm objectives (assessment of the achievement level, reasons for that and remarks on their advantages and others)
2. Results of the current status analyses of education and research activities performed by faculties, graduate schools, and institutes (levels of education and research and degree of improvement in their quality)
3. Objection and questions made by universities and responses given by the Committee for National University Education and Research Evaluation (if any observations and requests are made)

2. Evaluation of achievements against midterm objectives concerning education and research

For the midterm objectives of national university and inter-university corporations, NIAD-UE evaluated the attainment of "objectives concerning improvements in the quality of a university's education and research" and "objectives concerning improvements in the quality of a research institute's education and research", and also the contents of their midterm plans. To conduct comprehensive evaluations, NIAD-UE paid particular attention to whether the programs included in the midterm plans had produced outcomes, and whether the quality of education and research had improved or their high standards maintained during the midterm period, and also referred to the results of the current status analyses of faculties, graduate schools, and institutions.

In most cases, "objectives concerning improvements in the quality of a university's education and research," or part of national university and inter-university corporations' midterm objectives that NIAD-UE is to evaluate, contain what is shown in Table 3-1 (p. 59). Each university has its own structure of faculties, graduate schools, and institutions, and implements education and research activities suited to its own background, its location, and demands of society. To have a better understanding of each university, NIAD-UE asked them to describe their characteristics and distinctive characters, which was

expected to clarify the basic objectives included in the midterm objectives. When the evaluations of their achievements against their midterm objectives were conducted, those descriptions and the basic objectives were taken into account.

As for the achievement evaluation procedures, NIAD-UE divided the midterm objectives into three layers (major, middle and minor objectives), as seen in Table 3-1 (p. 59); it evaluated the minor objectives first, added up the results to evaluate middle objective, and eventually evaluated the major objectives. The evaluations were ranked for each major objective on a scale of five grades with reasons being stated. To carry out the evaluations, the results of the current status analyses of faculties, graduate schools, and institutions were referred to. Programs considered to be pioneering or ambitious were remarked upon, even if their targets had not been reached as planned, because NIAD-UE decided they were worth mentioning.

The results of the evaluation of national university corporations' and inter-university research institute corporations' achievements against their midterm objectives are shown in Table 3-3 (p. 66). No corporation was rated as "inadequate" or "major improvements needed," and many universities were rated as "adequate". The latter part can be attributed to the fact that in many cases the descriptions of the outcomes produced by their programs were insufficient or unclear. This can be seen in the verification survey conducted among the universities evaluated[14]; even though more than 70% of the universities responded that they described the outcomes, they also acknowledged that there were some midterm objectives and plans whose outcomes were difficult to describe. In the free descriptive answers, some universities pointed out that education and research hardly produce outcomes over a short period of time or requested that some examples be provided so they could tell what was appropriate to describe in the performance report, provisional outcomes such as improvements in how to teach, or final outcomes such as higher abilities of students.

Based on the findings of the verification work done through, among others, the questionnaire survey conducted among the universities evaluated and the external evaluators,[14] the evaluation of national university corporations' and inter-university research institute corporations' achievements against their midterm objectives can be summarized as follows:

Table 3-3 Results of the evaluation of achievements against midterm objectives
National university corporation (86 corporations)

	Objectives concerning education	Objectives concerning research	Other objectives
Excellent	0 (0%)	2 (2.3%)	2 (2.3%)
Good	9 (10.5%)	26 (30.2%)	32 (37.2%)
Adequate	77 (89.5%)	58 (67.4%)	52 (60.5%)
Inadequate	0 (0%)	0 (0%)	0 (0%)
Major improvements needed	0 (0%)	0 (0%)	0 (0%)

Inter-university research institute corporation (4 corporations)

	Objectives concerning research	Objectives concerning joint use	Objectives concerning education	Other objectives
Excellent	1 (25%)	0 (0%)	1 (25%)	0 (0%)
Good	1 (25%)	2 (50%)	1 (25%)	2 (50%)
Adequate	2 (50%)	2 (50%)	2 (50%)	2 (50%)
Inadequate	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Major improvements needed	0 (0%)	0 (0%)	0 (0%)	0 (0%)

Figures represent the number of corporations. Figures in parentheses represent the percentage of the total number. Other objectives mean the ones concerning partnership with society and international exchange.

The method of evaluation this time, meaning universities conducted a self-evaluation first and then the external evaluators of NIAD-UE carried out an evaluation, focused on the outcomes achieved under the midterm plans was largely favored. Most universities and external evaluators said that a major change in the way of evaluation that would nullify the experience gained through the evaluation work this time should be avoided. Thus it is important that the evaluation method be modified with objectives and plans for the second midterm period taken into account while maintaining the basic evaluation method this time. For example, it needs to be examined what should be regarded as outcomes achieved under the midterm plans or how uniform directions can be given to universities on supporting materials and data despite differences in the midterm objectives and plans among universities.

As for the effects and impacts of the evaluation this time, universities are

expected to be encouraged to find and rectify anything wrong in the PDCA (plan-do-check-act) cycle and to reflect on it in the formulation of the objectives and plans for the next midterm period. As for the effects and impact on education, many universities said that more and more attention was being paid to educational and learning outcomes and that the evaluation had contributed to an improvement in education on the whole. As for research, in contrast, even though the evaluation had an impact on systematic management, such as the formulation of strategies, it seems that the evaluation's overall contribution to improvement in research was somewhat smaller than improvement in education. This suggests that overall, the incorporation of national universities and inter-university research institutes is heading in the expected direction, but this direction has not been shared by all academic and administrative staff. Some respondents expressed concern that evaluation fatigue and the reality of evaluation work falling on the shoulders of a small number of academic staff could lead to formulating midterm objectives/plans in an overly easy way.

3. Current status analysis of faculties, graduate schools, and institutes

The current status of faculties, graduate schools, and institutes was grasped by analyzing the levels of their education and the degree of improvement in its quality, and the levels of their research and the degree of improvement in its quality. The levels of education and research indicate the situation of education and research activities and their outcomes at the time of evaluation. Thus, when the analyses were carried out, the purposes of education and research activities performed by the faculties, graduate schools, and institutes were taken into account. Purposes refer to the basic policies, directions, and basic outcome targets of their education and research activities.

The degree of improvement in the quality of education and research is measured by comparing and analyzing their levels at the time when national universities and inter-university research institutes were incorporated and at the time of evaluation. However, with the levels at the time of incorporation not necessarily clear, the degree of improvement was determined this time by analyzing process on how the universities' education and research activities and their outcomes had been improved and enhanced by the time of evaluation. And it is noteworthy that degree of improvement was assessed by taking into account of individual purposes of each university.

Levels of education

The levels of education were determined based on the current status of education activities performed by each faculty, graduate school, and institute, and their outcomes at the time of evaluation. So NIAD-UE adopted five analysis points (Table 3-4) and attached viewpoints to each of the analysis points. It also allowed faculties, graduate schools, and institutes to create their own viewpoints to suit their own purposes and situations. Some of them did so, but all of those viewpoints were in line with the original ones.

Table 3-4 Analysis points and their viewpoints on the standards of education

Analysis points	Viewpoints
I. Systems of teaching	○ Basic organizational arrangements ○ Systems aimed at improving what to teach and how to teach
II. What to teach	○ Arrangements for academic programs ○ Responses to requests from students and society
III. How to teach	○ Combined lesson styles and ingenious teaching methods ○ Efforts to encourage voluntary learning
IV. Learning outcomes	○ Knowledge, qualifications, and abilities acquired by students ○ Students'evaluation of their learning outcomes
V. Situation after graduation	○ Situation after graduation or completion ○ Evaluation by those concerned

The levels of education were assessed for each analysis point shown in Table 3-4 from the perspective of whether faculties, graduate schools, and institutes were meeting the expectations of the assumed stakeholders, after compiling the results of the analyses made from each of the above-mentioned viewpoints. Stakeholders mean people and organizations that directly or indirectly benefit from the education activities of the faculties, graduate schools, and institutes or their outcomes; in more specific terms, they are assumed to be students, examination applicants and their families, graduates, teachers and their employers, and local communities connected to the faculties, graduate schools, and institutes. The levels of education were ranked for each analysis point on a scale of four grades (Table 3-5) with the reasons being stated.

Let us summarize the findings of the verification work done through, among others, the questionnaire survey conducted among the universities evaluated and the external evaluators.[14] Sixty percent of the respondents, both universities and

Table 3-5 Results of the evaluation of the standards of education provided by faculties, graduate schools, and institutes

	Systems of teaching	What to teach	How to teach	Learning outcomes	Situation after graduation
Excellent	7 (0.9%)	6 (0.8%)	7 (0.9%)	6 (0.8%)	2 (0.3%)
Good	121 (15.1%)	157 (19.6%)	154 (19.2%)	85 (10.6%)	79 (10.1%)
Satisfactory	668 (83.4%)	636 (79.4%)	635 (79.3%)	687 (85.9%)	676 (86.8%)
Unsatisfactory	5 (0.6%)	2 (0.2%)	5 (0.6%)	22 (2.7%)	22 (2.8%)

The levels were assessed from the perspective of whether faculties, graduate schools, and institutes were meeting the explanations of the assumed stakeholders.
The total number of organizations evaluated was 801: 800 faculties, graduate schools and institutes were evaluated in terms of "learning outcomes" and 779 in terms of "situation after graduation" because some were not ready for evaluation. Figures represent the number of organizations. Figures in parentheses represent the percentage of the total number.

external evaluators, said they had little difficulty in describing or analyzing "I. Systems of teaching", "II. What to teach", and "III. How to teach". In contrast, only 40% to 50% of the respondents said they had little difficulty in describing or analyzing "IV. Learning outcomes" and "V. Situation after graduation". In other words, they found it more difficult to analyze the points concerning the output and outcomes of education than the points concerning the input and process of teaching. Analysis methods of learning outcomes will need to be developed and spread in due course. As for "IV. Learning outcomes", it is often pointed out that there is a limit to "student satisfaction questionnaires" because their results lack sufficient objectivity to reflect learning outcomes. Presumably, not only the results of the questionnaires but also how the situation has been improved based on them should be taken into account.

It was also found that both corporations and external evaluators wondered to some extent how to make assessments according to the criterion of "expectations from the assumed stakeholders". It is possible that the universities found it difficult to give a clear description on that point, and that the evaluators had difficulty in understanding what they had described. National university corporations are composed of a wide variety of organizations and each of them performs their own unique activities. Thus, the criterion of "expectations from the assumed stakeholders" itself is an essential point to make clear the individual character and roles of an organization. If the assumed stakeholders

are not clearly defined, it would be impossible to specify which organizations they would like to provide information, and also be difficult to form a consensus within the organizations.

The findings of the questionnaire survey showed that the implementation of current status analyses this time led academic staff to change their approach to performing activities and to start communicating better and sharing challenges with each other. Also, many respondents said that the current status analysis had given them an opportunity to reflect on their educational activities. Meanwhile, some respondents expressed concern over an adverse effect on activities that were difficult to evaluate properly. Many universities responded that they "will take a look in due course" at what to do with the evaluation results because the questionnaire survey was conducted soon after the results came to light; nonetheless, a large number of universities intend to make use of the evaluation results to, among other things, draw up objectives and plans for the next midterm period, reorganize themselves, enhance and improve their academic programs, and publicize their activities.

This was the first time for universities to compile current status reports on faculties, graduate schools, and institutes, and they had to take on a heavy workload; still, a large number of universities and external evaluators said it would be "possible to do the work more efficiently" if the same method continued to be used to conduct current status analyses. Also, many opinions were voiced that even though the uniqueness of each academic area should be recognized, creating a uniform format for the current status reports with regard to, for example, description items and description examples, and conducting evaluations based on definite indices were needed.

Levels of research

The levels of research were determined based on the current status of research activities performed by each faculty, graduate school, and institute, and their outcomes at the time of evaluation. NIAD-UE adopted two analysis points (Table 3-6) and attached viewpoints to each of the analysis points.

The levels of research were assessed for each analysis point from the perspective of whether faculties, graduate schools, and institutes were meeting the expectations of the assumed stakeholders, referring to the results of the

Table 3-6 Analysis points and their viewpoints on the standards of research

Analysis points	Viewpoints
I. Current status of research activities	○ How are research activities being performed? ○ For inter-university research institutes and universities' institutes and research centers with an inter-university research function, how are joint use/joint research being implemented?
II. Current status of research outcomes	○ Current status of research outcomes (In the case of inter-university research institutes and universities' institutes and research centers with an inter-university research function, the current status of the outcomes of joint use/joint research is to be included.)

analyses made from each of the above-mentioned viewpoints. Stakeholders mean people and organizations that directly or indirectly benefit from the research activities of the faculties, graduate schools, and institutes, or their outcomes. Those stakeholders are primarily assumed to be the relevant academic communities in terms of academic aspects, and the international community, local communities, and particular industries in terms of social, economic, and cultural aspects as well as the stakeholders mentioned in the "levels of education".

An analysis of the levels of research requires two analysis points: 1) current status of research activities, a point aimed at determining how the research activities are being performed by an entire organization based on numerical data, and 2) current status of research outcomes, a point aimed at determining the current status of the research outcomes produced by an entire organization using outstanding research performances representing the organization as supporting materials. As for selecting such research performances, each university submitted the ones regarded to be outstanding enough to represent the organization (research performances that would fall into the top two categories of the five grades based on the criteria shown in Table 3-7) from either an academic standpoint or social, economic, and cultural standpoints based on third-party evaluations, objective criteria, and other indices (Table 3-7). NIAD-UE evaluated the "current status of research outcomes" based on what peer reviewers in each academic area determined about the levels of the submitted research performances.

Table 3-7 Examples of "third-party evaluations and objective criteria/indices" used for selecting outstanding research performances representing an organization

Academic aspects (primarily assumed stakeholders: relevant academic communities)
- Publications in established journals in academic areas with refereeing.
- Book reviews and citations in specialized periodicals and newspapers.
- References in chronicles of research and papers on trends in academic communities, and references in the bibliographies of academic publications.
- The impact factors of specialized periodicals in which research performances were published, and the citation indices of papers in which they were referred to.
- Awards given by academic institutions at home and abroad.
- Established associations/international meetings in academic areas, at home and abroad, where guest lectures/ keynote speeches on research performances were delivered (names of associations/meetings and year when delivered should be mentioned).

Social, economic, and cultural aspects (primarily assumed stakeholders: international community, local communities, and particular industries)
- To what extent are research performances known/used, used/applied in local communities and particular industries, and actually reflected in policies?
- In the case of writing textbooks, books for the general public, and other publications concerning research areas, they are reviewed by respected book review publications or they have long been widely used and influential.
- In the case of artistic performances, their outcomes are highly praised by respected critics in the fields.

Note: Social, economic, and cultural contributions mean that the research performances themselves are helping in specific ways in terms of social, economic, and cultural aspects. So even if the academic staff concerned participate in social activities (for example, being members of councils or other bodies of central or local governments), that participation would not constitute evidence here.

Table 3-8 Results of the evaluation of the standards of research conducted by faculties, graduate schools, and institutes

	Current status of research activities	Current status of research outcomes
Excellent	34 (5.5%)	26 (4.2%)
Good	248 (40.4%)	231 (37.6%)
Satisfactory	327 (53.3%)	354 (57.7%)
Unsatisfactory	5 (0.8%)	3 (0.5%)

The levels were assessed from the perspective of whether faculties, graduate schools, and institutes were meeting the explanations of the assumed stakeholders.
The total number of organizations evaluated was 614. Figures in parentheses represent the percentage of the total number.

The levels of research were ranked for each analysis point on a scale of four grades (Table 3-8) with reasons being stated. Let us summarize the findings

of the verification work done through, among others, the questionnaire survey conducted among the universities evaluated and the external evaluators.[14]

As for selecting the outstanding research performances representing the organization, only about 30% of the universities responded they "were able to make definite decisions" on the level of their academic significance. Many respondents acknowledged that different criteria should be applied to different academic areas; still, a large number of respondents suggested that NIAD-UE establish certain more specific criteria. It was pointed out that for the faculties of education, engineering, and interdisciplinary studies in particular, which all cover a wide range of research areas, it was difficult to make decisions on the level of academic significance within the organization. Meanwhile, more than half of the external evaluators responded they had been able to make definite decisions, and only about 10% of the external evaluators said they "often had difficulty in making decisions".

As for evidence for academic significance, the number of citations and impact factors tended to be commonly used as criteria because it was difficult to find evidence other than these. Also, it was pointed out that it was hard to find appropriate evidence in the areas of the humanities and social sciences. Some of the universities were in favor of only outstanding research performances being used as supporting materials for analysis. In addition, more than 70% of the respondents said "only explanation based on evidence for academic significance should be provided and the papers themselves should not"; thus it is safe to say that the method this time was largely approved of. Nonetheless, as some of the respondents in the free descriptive answers said that "since outstanding research performances that represent an organization should be widely known, only naming such performances would be enough for peer reviewers to make a judgment", it needs to be considered to what extent universities should be asked to describe evidence used in peer reviews to determine the levels of research performance.

As for social, economic, and cultural significance, it seemed that for both universities and external evaluators it was difficult to give evidence and make judgments, regardless of the academic area. Consideration will be needed in due course on the examples of evidence and the ways of making judgments.

With university evaluation still in its infancy, a great amount of work was done this time; still, given that many respondents said it would be possible to do the work more efficiently with the same framework used, it is reasonable

to assume that major changes in the method of evaluation should be avoided. The findings of the questionnaire survey revealed a large number of opinions respecting the uniqueness of each academic area, but using descriptive items suited to each academic area and giving descriptive examples were needed in the current status reports. As for the explanatory sheet on research performances, due consideration is needed with the following taken into account, as pointed out in the questionnaire survey: in some academic areas it is hard to present supporting materials; it is difficult to provide evidence to enable judgments on social, economic, and cultural significance; and it is difficult to make judgments within a university because of the interdisciplinary structure of faculties, graduate schools, and institutes. Now that the national university corporation evaluation for the first midterm period is finished, an exchange of information among universities, and between universities and NIAD-UE, on what has been experienced is called for.

Notes:

1. OECD (1997). In Search of Results: Performance Management Practices.
2. Osborne, D., Gaebler, T. (1992). Reinventing Government: How the Entrepreneurial Spirit is Transforming the Public Sector, Reading, Mass: Addison Wesley.
3. Pollitt, C., Bouchkaert, G. (2000). Public Management Reform, Oxford University Press.
4. Stigler, G.J. (1971). "The Theory of Economic Regulation", Bell Journal of Economics and Management Science, no. 3, pp. 3-18.
5. The basic plan for downsizing national administrative organizations and improving their efficiency, decided by the cabinet in April, 1999. Available on the office of the prime minister's website at: http://www.kantei.go.jp/jp/kakugikettei/990524sosiki.html (only in Japanese)
6. "Structural Reform of the Japanese Economy: Basic Policies for Macroeconomic Management," decided by the cabinet in June, 2001. Available on the office of the prime minister's website at: http://www.kantei.go.jp/jp/kakugikettei/2001/honebuto/0626keizaizaisei-ka.html (in Japanese)

7. The report concerning the image of national university corporations, released by a research council of the Ministry of Education, Culture, Sports, Science and Technology (MEXT) in March, 2002. Available on MEXT's website at: http://www.mext.go.jp/b_menu/shingi/gijyutu/gijyutu8/toushin/0211021.html (in Japanese)
8. "Basic Policies for Economic and Fiscal Policy Management and Structural Reform 2002," decided by the cabinet in June, 2002. Available on the office of the prime minister's website at: http://www.kantei.go.jp/jp/singi/keizai/kakugi/020625f.html (in Japanese)
9. "Guidelines for Formulation of the FY2003 Budget," decided by the Cabinet in November, 2002. Available on the office of the prime minister's website at: http://www.kantei.go.jp/jp/kakugikettei/2002/1129yosan.html (in Japanese)
10. The outline of the national university corporation system compiled by MEXT. Available on MEXT's website at: http://www.mext.go.jp/a_menu/koutou/houjin/03090201.pdf (in Japanese)
11. Hatry, H.P. (1999). Performance Measurement: Getting Results, The Urban Institute Press
12. Evaluation of education and research activities of national universities and inter-university research institutes. Available on NIAD-UE's website at: http://www.niad.ac.jp/n_hyouka/kokuritsu/1179353_926.html (in Japanese)
13. The results of the evaluations of education and research activities of national university corporations and inter-university research institute corporations during the first midterm period. Available on NIAD-UE's website at: http://www.niad.ac.jp/n_hyouka/kokuritsu/kekka_h22/index.html (in Japanese)
14. Verification reports on the evaluations of education and research activities of national university corporations and inter-university research institute corporations. Available on NIAD-UE's website at: http://www.niad.ac.jp/ICSFiles/afieldfile/2010/02/26/no6_3_houkokusyo.pdf (in Japanese)

Afterword

At the commencement of its trial university evaluation, the National Institution for Academic Degrees and University Evaluation (NIAD-UE) coined and started to use the phrase “evaluation culture.” Much to our pleasure, the coined term has recently been put into remarkably widespread use; however, we have come to realize the enormousness of our responsibility. During the period of the trial university evaluation, NIAD-UE aimed at fostering the evaluation culture; as NIAD-UE embarked on its certified evaluation and accreditation of higher education institutions and its performance-based evaluation of national university corporations, the institution intended to further develop the culture; and now it is safe to say that the culture is taking root, meaning that higher education institutions are striving to improve and enhance the quality of education and research with the strategic use of the results of the evaluations. This can be proved by the fact that the government’s initiative to reform the higher education system has been undertaken based on evaluation culture. So our next hope is that the culture will enter the maturity stage in which the stakeholders find the results of the evaluations useful and take full advantage of them.

The trial university evaluation started in 2000. Since then, certified evaluation and accreditation of higher education institutions and professional graduate schools have been implemented, along with evaluation of national university corporations’ performance in a six-year period aimed at attaining their midterm objectives for education and research. So it can be said that universities and evaluation organizations have dedicated themselves to working on their own evaluation tasks. But we have to admit that the intended functions of university evaluation have yet to be fully understood by those concerned, such as universities, third-party evaluation bodies and maybe even society. Thus, evaluation fatigue is becoming a serious problem. A cause of the problem could be that evaluators usually have to carry out more than one task, inevitably bringing about repetition and redundancy. From this angle too, they need to clearly understand the three functions of evaluation: accreditation, audit and assessment.

As the world rapidly becomes borderless and more connected, questions are being raised about whether Japanese universities will be able to lead the world's academic community, and whether Japanese youths will be able to play a leading role in the international arena. In addition, given that Japanese students are generally said to be subpar in terms of academic achievement, there is doubt as to whether they are smart and tough enough to survive in today's globalized world. There is also the question of whether they have the skills to communicate with people whose backgrounds are different from their own. No country can win hearts and minds with superior economic or military power alone anymore. In the globalized 21st century, individuals and countries have to be humble enough to understand and respect each other's differences.

Universities around the world are working hard to go global with a sense of urgency. Nonetheless, drawing students from foreign countries is not their primary goal; what they really intend is to become global institutions, thereby helping to improve and enhance the quality of the education they provide, to nurture bright people who can contribute to the international community. So globalization does not mean applying a set of universal standards to everything, as each country, region and university has its own established customs and traditions; it means accepting something different from what they are familiar with while preserving their own customs and traditions in hope of creating synergies. This is why quality assurance based on both the entity's individual character and standard features matters.

Finally, we gratefully acknowledge the staff of the Research Department and the Department of University Evaluation of NIAD-UE, and many people outside NIAD-UE. They provided great help when we were preparing to launch this latest publication in NIAD-UE's University Evaluation series. We would also like to express my gratitude to the members of the Board of Councilors, the members of the Administrative Committee and evaluators of NIAD-UE, who always shared with me their interesting and insightful thoughts.

Authors:

Akihiko KAWAGUCHI Foreword, Chapter 2, Chapter 3 and Afterword
(Specially Appointed Professor, NIAD-UE)

Yayoi TANAKA Chapter 1 and Chapter 3
(Associate Professor, Research Department, NIAD-UE)

National Institution for Academic Degrees
and University Evaluation

1-29-1 Gakuen-nishimachi
Kodaira, Tokyo 187-8587 Japan

Tel *+81 42 307 1616*
Fax *+81 42 307 1559*
Email *kokusai@niad.ac.jp*
http://www.niad.ac.jp