

# 大学における教育研究活動の評価に関する調査研究

文部科学省平成23年度先導的大学改革推進委託事業

研究成果報告書

平成23年（2012年）3月

研究代表者 北原 和夫

東京理科大学大学院科学教育研究科 教授



## まえがき

北原和夫（東京理科大学）

この報告書は、平成 22 年度先導的大学改革推進委託事業として文部科学省から委託された調査「大学における教育研究活動の評価に関する調査研究」の成果をとりまとめたものである。

2004 年の認証評価の制度化及び国立大学法人制度の発足によって、わが国の大学における教育研究活動は、評価のシステムとの新しい関係を持つことになった。その後の数年間に、認証評価と国立大学法人評価の実践が積み重ねられてきて、両方の制度がともに 2 期目のサイクルにさしかかっている。これら二つのほかにも、国際ランキングや質保証の議論などにあらわれているように、多様な評価の仕組みが議論の俎上に据えられ、各種の取り組みが議論されたり、試みられたりしている。1990 年代が評価の必要性が叫ばれ、制度導入をめぐる動きが始まった 10 年だとすると、2000 年代初頭の 10 年は、評価制度がある程度制度として導入・定着してきた 10 年だといえる。

制度の導入・定着が進行したこの 10 年ほどの間に、現実の大学評価システムの成果だけでなく問題点や難しさも次第に明らかになってきているように思われる。うまくいっているようにみえる部分がある一方で、思いがけない問題が出てきているような部分も見受けられる。わが国の大学がある程度評価に慣れてきた現在、ルーティーンで今の評価の仕組みを動かすことに安住せず、もう一度原理や理念に立ち戻って、大学における教育研究活動の評価のあり方をとらえ直すことが必要なのではないか。われわれの研究は、そうした問題意識でスタートした。

大学にとって教育研究の評価とはいったいどういう意味があるのか、評価の仕組みによって、大学はどういう影響を受けているのか、現状の評価の仕組みに問題があるとするそれは何なのか、これからの大学にとって（そして社会にとって）有益な教育研究の評価の姿はどのようなものか——そのような問題を考えるために、われわれの研究は、3 つのグループに分かれて次のような検討を行った。

第 1 グループでは、「大学における教育研究活動の評価のシステムに関する国内外の動向の整理」を目的に掲げて検討を行った。欧米における評価の動向を整理しつつ、今後の望ましい評価のあり方について議論するための論点をさぐるとともに、日本の評価制度については、特に認証評価のあり方に焦点を絞って議論を行った。認証評価の仕組みの中に教育研究の質の問題を位置づけるときの重要な課題の一つが、分野別の質保証の問題である。これについては、現在日本学術会議が策定を進めている分野別の教育課程編成上の参照基準が、認証評価の枠組みの中にどのように位置付けられたらよいのかについて、集中的な議論を行った。

第 2 グループでは、「大学に対するパブリックマネジメントの在り方に関する調査研究」を目的に掲げて検討を行った。大学は（私立大学も含めて）公的な資金が投入されている機関であり、必要な業務目標を達成し、その状況を国民に開示する責務を負っている。しかし大学の持つ機能の多元性や、教育や研究が必要とする長期的な視点は、目標設定型の管理手法になじまない面も有している。また、日本の大学が置かれてきた状況に由来する固有の文脈も考える必要がある。ここでは、国立大学法人評価について、その制度的背景や枠組みについて、比較的自由的な観点から見直してみた。

第3グループでは、「今後の日本社会とそこでの大学の在り方に関する調査研究」を目的に掲げて検討を行った。大学は社会のあり方に規定される存在であると同時に、新しい社会を生み出していく知的源泉でもある。経済や政治や文化が大きく変容しつつある状況の下では、大学は社会の変化を見通して自らのあり方を考える必要がある。大学における教育研究の評価のあり方もまた、社会の変化の中での新しい大学像に適合的なものでなければならぬ。そこで、今後のグローバルな世界のありよう、日本社会の変容のありようについて、自由闊達に議論しつつ、そこで求められる大学のあり方と教育研究の評価のあり方について、意見を交換し、広く共有しうる見解をとりまとめた。

このように、われわれの検討した課題は多岐にわたっている。それらをふまえて、序論では、評価をめぐる基本的な考え方を原理レベルで考察し、その後、第1グループ～第3グループの検討の成果を、第1部～第3部の柱立ててとりまとめた。

全体の検討の中で浮かび上がってきたように思われるのは、「あるべき大学像は1つではないだろう」ということと、「それぞれの大学が自らの公共的使命をどう考えるのかを真剣に考える必要があるだろう」ということである。大学における教育研究の評価のあり方を考える我々の研究課題に即して言うところ、評価のあるべき方向は、その二つのポイントを支持し、促進するものでなければならないだろうということである。

「あるべき大学像は1つではないだろう」ということは、近年盛んに論じられている「大学の機能別分化」論と関わっている。どの大学もただ一つの基準やただ一つの規範に向けて組織されるとしたら、それはかえって多様な大学が存在することによって生まれる教育の豊饒さや知の広がりを阻害してしまうことになるだろう。大学は多様である必要がある。

しかしながら、同時に、大学を機能別に種別化・固定化することをどこまで進めるべきかは微妙な問題をはらんでいる。どの大学であれ、大学は新しい目標に向けて自らの姿を作りかえていく流動性をもつべきだし、多様な社会的ニーズや知的要請に応える複合性をもつべきである。また、一つの大学内にさまざまなニーズを持った学生がいることもまた、大学に複合的な機能を求めている。だから、各大学の機能を外部から限定づけたり固定化させたりしないことこそが、大学の潜在的な力の発揮に重要ではないかと考えることができる。そうした大学の多様性を生かすような評価の仕組みが求められよう。

とはいえ、目の前にあるような大学の多様性がそのままよいともいえないだろう。欧米でも日本でも、大学という組織は多元的で分散的であることがその組織上の特徴である。それゆえ、下手をすれば既得権とルーティーンのみが支配する場となってしまうケースもありうる。

ここで、先ほど述べた「それぞれの大学が自らの公共的使命をどう考えるのかを真剣に考える必要があるだろう」という点が意味を持つ。近年は「ステークホルダーを意識した大学」が強調されているが、われわれが考えるに、声が大きくまた直接的な「見返り」を求める目の前のステークホルダーが、大学にとって考慮すべきすべてではない。第1に、教育効果の長期性・外部性があるから、大学は未来の社会・世界全体にとってどういう意味を持つのか問われる。第2に、グローバル化が進む世界における大学では、教育研究の成果が世界全体とつながっていることが考慮されなければならない。それゆえ、目の前の学生の声から未来世代への責任、日本国内のみならず世界の国々の様々な立場の人たちへの責任まで、われわれは、幅広いステークホルダーを意識して大学の果たすべき役割について考えてみる必要がある。もちろん、幅

広いステークホルダーのどの部分に注目して何をやろうとするのかは、各大学（あるいは最終的には各大学人）の選択に委ねられている。しかし、誰をステークホルダーに選んだとしても、その選択は幅広いステークホルダーが存在する中でなされるものであるかぎり、その選択が持つ公共性はきちんと社会に対して説明されるべきものでなければならないだろう。

そのように考えると、大学評価における説明責任（Accountability）とは、「社会一般」への顧客サービスではなく、公共性を踏まえて自覚的に選択したステークホルダーに対して自らの実績を説明することである。多くの危機や課題を抱える 21 世紀は、人類が協働して危機や課題に取り組まねばならない時代である。各大学は、教育研究という知的な営みを共通項としてもちながら、それぞれの大学が設立理念、人的物的資源等にとって最適な公共的使命を掲げ、現在の世代や未来の世代にとってより有益な存在になっていくように不断の努力を続けていくべきである。そして、大学における教育研究の評価は、最終的にそのような不断の努力を支援するものでなければならない。

# 目次

まえがき

目次

序論 教育研究の評価をどう考えるか … 1

北原和夫(東京理科大学)・広田照幸(日本大学)

第1部 大学における教育研究活動の評価のシステムに関する国内外の動向の整理 … 33

第1章 評価に関する議論の整理と今後の課題 … 35

広田照幸(日本大学)・二宮祐(一橋大学)

第2章 米国のアクレディテーション——日本の認証評価との比較の視点から—— … 57

福留東士(広島大学)

第3章 欧州の評価の動向 … 69

大場淳(広島大学)

第4章 認証評価制度の問題点と第2期目の課題 … 83

瀧澤博三(日本私立大学協会附属私学高等教育研究所)

第5章 大学の認証評価と専門分野別質保証 … 91

早田幸政(大阪大学)

第6章 大学教育の質保証と評価の在り方について … 111

藤田英典(共栄大学)

第2部 大学に対するパブリックマネジメントの在り方に関する調査研究 …133

はじめに 第2部会の目的と実施スケジュール・論点と提言 …135

田中弥生(大学評価・学位授与機構)・丸山和昭(福島大学)

第1章 国立大学法人化の政策過程 …141

丸山和昭(福島大学)・田中弥生(大学評価・学位授与機構)

第2章 国立大学法人制度の骨格と運営の実際 …167

田中弥生(大学評価・学位授与機構)・丸山和昭(福島大学)

|                                      |                                   |             |
|--------------------------------------|-----------------------------------|-------------|
| 第3章 第2部会の議論集                         | 第2部会の議論の要約                        | …191        |
|                                      | 田中弥生（大学評価・学位授与機構）・丸山和昭（福島大学）      |             |
|                                      | 第1回検討会の議論                         |             |
|                                      | 第2回検討会の議論                         |             |
|                                      | 第3回および第4回合同検討会の議論                 |             |
|                                      | 第5回検討会の議論                         |             |
| <b>第3部</b>                           | <b>今後の日本社会とそこでの大学の在り方に関する調査研究</b> | <b>…325</b> |
| はじめに                                 | 日本の大学は生き延びられるか 第三部主査としての小括        | …327        |
|                                      | 高祖敏明（上智大学）                        |             |
| 第1章                                  | 今後の社会と大学の在り方                      | …331        |
|                                      | 白川優治（千葉大学）・崎山直樹（千葉大学）             |             |
| 第2章                                  | 第3部会の議論の要約                        | …341        |
|                                      | 崎山直樹（千葉大学）                        |             |
|                                      | 第1回検討会の議論                         |             |
|                                      | 第2回検討会の議論                         |             |
|                                      | 第3回検討会の議論                         |             |
|                                      | 第4回検討会の議論                         |             |
| 講演資料1                                | 大学モデルの再検討？ 知識の創造場所としての大学と普遍性      | …349        |
|                                      | 上山隆大（上智大学）                        |             |
| 講演資料2                                | 日本の大学にとっての国際化                     | …353        |
|                                      | 北村友人（上智大学）                        |             |
| <b>第4部</b>                           | <b>研究のまとめ</b>                     | <b>…379</b> |
| シンポジウム記録「パネル討論：大学における教育研究活動の評価をめぐって」 |                                   | …381        |
| 研究の総括                                | 北原和夫（東京理科大学）                      | …413        |
| 研究組織と活動の記録                           |                                   | …416        |
| 奥付                                   |                                   |             |

## 序論 教育研究の評価をどう考えるか

北原 和夫（東京理科大学）・広田 照幸（日本大学）

### はじめに——本章のねらい

大きな社会変動が生じている現代日本社会において、その変動の影響を受けつつ、この20年間ほどの間に日本の大学はさまざまな改革の試みが積み重ねられてきた。その大学改革の中で常に重要な焦点となり続けてきたのが、大学における教育研究活動の評価のあり方の問題である。本委託調査では、この問題をどう考えたらよいのかについて検討する。この章は、その総論として、大学の教育研究の評価をどう考えるかについての概括的な考察を行っておきたい。

欧米では、1980年代から大学における教育研究の「質」(quality)への関心が急速に高まった。米国では、1980年代前半から、高等教育の質への関心が高まり、多様な学内情報が集められ、吟味にかけられる仕組みが急速に普及し、それを正当化する理論化も進んだ。大部分の高等教育機関ではフォーマルな機関レベルの評価の活動が発展した。その背景には、①高等教育拡大が終わり、各機関が財政支援を求める新たな理由として「質」を掲げるようになったこと、②初中等教育の危機が問題とされ、高等教育まで含めて学校改革への投資の必要性論が浮上したこと、③経済発展のエンジンとして戦略的投資の対象としてみられるようになったこと、といった社会の変化が存在していた (Ewell 2007)。そうした教育研究の「質」への関心は、1970年代までにすでに長い歴史をもって定着していた米国の認証評価の仕組みに、新たな社会的・政治的要求を課していった。新しい評価尺度や評価手法の登場や、評価の利用に関する新しい仕組みの導入などである。

欧州諸国でも1990年代には、評価を介した高等教育の質保証の動きが進んだ。社会主義体制が崩壊した東欧では、国家の管理・統制を離れた大学が市場経済の中で自らの「質」を保証するために、評価システムを活用する取り組みが活発化した。西欧諸国では、急速な大学進学率の増加や知識経済化への国家による関心の高まりなどによって、評価機関の設置や評価システムの導入が進んでいった (Schwarz and Westerheijden 2007)。羽田貴史 (2004) は、高等教育におけるガバナンスのアクター (政府—同僚—市場) と、保証される対象 (研究—教育—経営) との2つの軸でさまざまな質保証の制度を整理している (表1)。実にたくさんの種類の制度が作られてきていることがわかる。



| アクター<br>対象 |                                               |                        | 政府による規制                                              | 専門的権成による同僚規制<br>職能代表・養成機関・<br>教育行政・学者      | 市場による規制                         |
|------------|-----------------------------------------------|------------------------|------------------------------------------------------|--------------------------------------------|---------------------------------|
| 研究         |                                               |                        | 研究費配分(RAE)<br>(UK)                                   | 同僚・専門家による評判<br>学会活動<br>学会の倫理コード            | 研究費の競争による配分<br>出版・ジャーナル<br>学会活動 |
| 教育         | へ機<br>の関<br>規・<br>制・<br>プ<br>ロ<br>グ<br>ラ<br>ム | 設置認可<br>charter        | 大学設置認可<br>(JAPAN)                                    | 課程認可審査(UK,AUS)                             | —                               |
|            |                                               | 認証評価<br>accreditation  | —                                                    | 評価認証(US)<br>大学基準協会(JAPAN)<br>専門プログラム(ABET) | —                               |
|            |                                               | 査定<br>assessment       | HEFCE(UK)                                            | —                                          | —                               |
|            |                                               | 視察<br>audit            | 行政監察(JAPAN)                                          | AUQA(AUS)                                  | —                               |
|            | 出<br>口                                        | 大学卒業水準                 | 医師試験<br>教員採用試験                                       | 外部試験員制度(UK)                                | 雇用者                             |
|            | 入<br>口                                        | 大学入試・<br>中等教育の<br>卒業水準 | センター試験(JAPAN)<br>GCE(UK)<br>アビトゥア(GER)<br>バカロレア(FRA) | 受験資格判定<br>入学者選抜・判定                         | 学生の学校選択                         |
| 経営         |                                               |                        | 会計監査                                                 | 自己点検                                       | 学生の学校選択を<br>介した資金・授業料配分         |

出典：羽田（2005）

表 1 質保証の諸制度

| いわゆる「第三者評価」*                                 |                                                   |                      |                                  |                                 |                                      |                       |
|----------------------------------------------|---------------------------------------------------|----------------------|----------------------------------|---------------------------------|--------------------------------------|-----------------------|
| 大学種別                                         | 学校教育法に基づく自己点検・評価及び認証評価                            |                      |                                  |                                 | 国立大学法人法<br>に基づく国立大<br>学法人評価          | その他の第三者<br>評価（JABEE等） |
|                                              | 自己点検・評価<br>外部評価                                   | 認証評価                 | 専門職大学院の認証評価                      |                                 |                                      |                       |
|                                              |                                                   |                      |                                  | うち法科大学院<br>の適格認定                |                                      |                       |
| 国立大学                                         | 必須、時期の定め無し<br>（当該大学の教育・研究・組織・運営・施設・設備）・外部評価の努力義務化 | 必須<br>7年に1度<br>（機関別） | 専門職大学院を持つ大学は必須<br>5年に1度<br>（分野別） | 法科大学院を持つ大学は必須<br>5年に1度<br>（分野別） | 必須<br>6年に1度<br>（各法人が定めた中期目標・中期計画に基づく | 任意                    |
| 公私立大学                                        |                                                   |                      |                                  |                                 | *公立大学法人は、法人評価が必要<br>*私立大学については不要     |                       |
| *いつれの第三者評価も被評価者である大学の提出報告書に基づいたピアレビューを基本とする。 |                                                   |                      |                                  |                                 |                                      |                       |

出典：田中（2011）

表 2 大学種別に見た評価の制度

日本でも、周知の通り、1986年の臨時教育審議会第二次答申での提起を受け、1991年に各大学の自己点検・評価が求められることになり（1999年に義務化）、さらに第三者評価の必要性が論じられ、2004年から認証評価制度がスタートした。また、同年には国立大学法人制度も発足し、その国立大学法人評価の中で教育研究の評価にあたる部分を大学評価・学位授与機構が行う仕組みがスタートした。田中弥生と丸山和昭は、それら二つが中心的な制度となる日本の評価制度の現状を表2のように整理している（田中2011）。

大学の教育研究の評価がある程度定着してきた現在、その制度の功罪をきちんと吟味し、必要な形で改善・改良を進めていくことが必要な段階にきているように思われる。大学の

評価に関して用いられてきているさまざまな評価の特徴や問題点・限界などを、かなり広い視点で整理し、具体的な課題を論じるのは別の章で行うことにして、ここでは、総論的に、大学における教育研究の評価が求められるようになった文脈と、そこで必要な改革の方向について、簡単に論じてみたい。

## 第1節 評価が不可欠になってきた文脈

### (1) グローバル化

まずは、日本の高等教育システムに関して評価の問題が浮上してきた社会的文脈として、ここでは2つの点を確認しておきたい。

第一に、いうまでもなく、グローバル化の影響を挙げることができる。物や人やカネや情報が簡単に国境を越えて移動するようになってきたとき、それぞれの国の高等教育は、いやおうなしにグローバルなレベルでの競争や連携・協働と関わりを持つことになる。これまでの大学評価の議論においても、グローバル化に対応する高等教育のあり方の見直しという文脈が、くり返し語られてきた。たとえば、学位や資格は今や一国内にとどまらず、グローバルなレベルで質を保証されたものであることが求められるようになった。留学生の受け入れや送り出し、卒業生のグローバル労働市場での評価などに関わり、国境を越えた次元での大学教育の質保証が問題となることになった。あるいは、グローバルに活躍する高度な人材をいかにして養成するのかといった、より教育内在的な課題も浮上しており、それが評価のあり方の議論と関わっている。さらに、大学が行う研究の次元においても、データベースの普及や評価指標の開発などにより、グローバルなレベルでの評価がさまざまな形でなされるようになってきている。こうしたことの詳細は、すでにたくさんの論文や本が書かれているので、ここで繰り返すまでもない。

とはいえ、グローバル化と高等教育の評価との関係について、あまり議論されてきていない重要な点が2つあると思うので、ここで少し書いておきたい。

一つは、グローバル化のインパクトを、国家間の経済競争の文脈だけで考えてしまわないことの重要性である。大学の教育研究が果たすべき役割は、経済に資する人材の育成や経済成長のための技術開発などにとどまるわけではない。このことは、中央教育審議会の答申でも、十分意識されてきている。

学部段階教育のあり方の方向をとりまとめた2008年12月24日の中教審答申「学士課程教育の構築に向けて」では、次のように書かれている。

グローバル化する知識基盤社会、学習社会にあつては、国民の強い進学需要に応えつつ、国際的通用性を備えた、質の高い教育を行うことが必要である。国境を越えた多様で複雑な課題に直面する現代社会にあつて、大学として、自立した21世紀型市民を幅広く育成することは、個人の幸福と社会全体の発展それぞれの観点で極めて重要であり、

公共的使命と言える。先進諸国の大学では、自らの使命を、学生の身に付ける学習成果という形で明示し、その達成度を評価するなどの取組が広がりつつある。

ここで「自立した 21 世紀型市民」とは、「専攻分野についての専門性を有するだけでなく、幅広い教養を身に付け、高い公共性・倫理性を保持しつつ、時代の変化に合わせて積極的に社会を支え、あるいは社会を改善していく資質を有する人材を指」しており、2005 年 1 月 28 日の中教審答申「我が国の高等教育の将来像」の中で出てきたものである。2008 年答申では、大学の使命として、「国境を越えた多様で複雑な課題に直面する現代社会」において、「社会を支え、……改善していく資質を有する人材」を育成することを掲げているのである。「学習成果の評価」もその記述の延長上に位置づけられており、大学の公共的使命に関する十分な自覚が大学に求められていることを読みとる必要がある。

また、大学院教育を扱った 2011 年 1 月 31 日の中教審答申「グローバル化社会の大学院教育―世界の多様な分野で大学院修了者が活躍するために」においては、次のように書かれている。

知識基盤型社会が進展し、知識・情報・技術の創造と活用が社会のあらゆる領域での活動の基盤として重要性を増す中、国内外の社会の様々な分野で活躍できる高度な人材が求められている。特に、世界が優れた知恵で競い合う時代に、専門分化した膨大な知識の体系を俯瞰しながら、イノベーションにより社会に新たな価値を創造し、人類社会が直面する課題を解決に導くために、国際社会でリーダーシップを発揮する高度な人材が不可欠となっている。

この答申でもまた、グローバル化に対応した大学院教育のあるべき方向は、「人類社会が直面する課題を解決に導くため」という使命に向けられていることに留意すべきである。

大学評価をめぐる議論の中では、われわれはつい、厳しい経済競争の中でのわが国の生き残りや、自大学の卒業生の良好な就職などにのみ関心が向きがちである。しかしながら、大学が果たすべき使命は、そうしたレベルのものにとどまらない。日本の大学が教育研究を通して実現していかないといけないのは、「国境を越えた多様で複雑な課題」や「人類社会が直面する課題」の解決に資する人材であるといえよう（もちろん、経済的な発展をもたらす人材はその中に含まれるが）。

もう一つ、グローバル化と高等教育の評価との関係について留意すべきは、評価システムのグローバル化をどう考えたらよいかという問題である。言い換えると、評価の制度や技法、尺度などは、どこまでグローバルに一元化されるのか、という問題である。評価の仕組みや技法は国境を越えていやおうなしに一元化される側面もあるし（同一尺度での比較など）、ある国の特定の仕組みや技法が借用や模倣、購入といった形で別の国に移植される場合もある。それとは別に、それぞれの国が固有の特徴を持った質保証の仕組みや評価

の仕組みをもつこともありうる。何でもかんでも他国のものを輸入して制度化すればよいわけではなく、それは前節で述べた「評価の氾濫」をもたらしてしまうことになる。近年進行しているのは、それぞれの社会に適した評価の仕組みを作り、比較可能な枠組みとして説明できるようになることを目指す、という動きである。われわれも、その線で今後の方向を考えていかねばならないだろう。

## （２）大学の大衆化

もう一つ、大学教育の評価が問題化してきた重要な要因は、大学進学率の上昇による大衆化である。大学進学率が同世代の５割を超えるようになってくると、旧来の大学生像、大学教育像にはおさまらない多様な学生や大学教育が散見されるようになってきた。「フランク大学」、「ボーダーフリー大学」などの語が作られ、それらの大学での学生や教育が、おもしろおかしく、あるいはため息混じりに語られるようになるにつれて、「大学生として十分な学力や意欲をもっているのか」とか、「大学教育の名にふさわしい教育をやっているのか」といった疑念が、社会から大学に対して浴びせられることになった。

中教審学士課程答申では、「ユニバーサル化」という語で、この事態を表現している。いうまでもなく、M・トロウが提出した高等教育拡大モデルの中で用いられた語である（トロウ 1976）。同答申では、現在の大学進学率等を過剰とする見方を斥け、「本審議会としては、若年人口の過半数が高等教育を受ける現状を積極的に受け入れていくこと、少なくとも、成績中位層以上の高校生が経済的理由により進学を断念せざるを得ない状況を無くしていくことが必要である」という立場を取っている。

教育機関としての大学が多すぎるのか、そうでないのかは、大学や学位という社会的制度の本質や機能をどのように考えるか、また、短期的・長期的に大学教育が生む成果をどう考えるかなどによって、意見が分かれるであろう。しかし、社会からの疑念がある以上は、大学の名にふさわしい教育とはどういうものであるかを明確にしつつ、それを実際に提供していることを、大学は社会に向かって責任をもつ必要がある。「大学とは何か」についての国家による明確な定義や基準が設定されず、多様な高等教育機関が叢生した米国で早くからア kreditation) のシステムが発達したのはそうした理由からであるし、旧来にないタイプの大学生や大学教育がみられるようになった日本でも、同様の課題に応えるべきことが求められる状況になっている。2004年に制度化された日本の認証評価制度は、そのような意味をもっているといえることができる。

## 第２節 評価のあり方を見直す視点

いずれの国や地域においても、大学における教育研究の評価の制度がさまざまに作られ、作動してきた。しかしながら、どこの国や地域でも、評価の制度化や実施をめぐるのは多かれ少なかれ、多様な種類の不満があふれている。

大学の第一線の教員からは、評価に伴う作業の繁雑さや評価項目・評価尺度に振り回された改革への不満が表明されている。たとえば英国では、ピア・レビューによる評価の負担の過重への不満と尺度化による評価の部分性のもたらす不満とに挟撃されて、評価の在り方が揺れ動いているようにみえる（本報告書大場論文を参照）。また、長く国家による強い統制のもとにおかれてきた東欧諸国を除くと、大学評価が大学にとって必要な自治や自律性を掘り崩してしまうものだという批判も出されている。もう一方で、大学外の人たち——たとえば産業界の人たちや教育行政官僚——からは、導入された評価制度が形骸化しているとか、社会に向かって発信される評価結果の情報がまだ不十分だという不満がしきりに出されている。

そもそも、評価をめぐる議論自体が混乱している部分がある。ヨーロッパでは、90年代初めには、「質」の定義もはっきりせず、さまざまな権限（*accreditation, evaluation, audit, benchmarking*、機関レベル／プログラムレベル）を持った国家レベルの質保証機関が、雨後の筍のごとく作られた。そこでは、内部質保証と外部評価とのバランスの関係がきちんと配慮されてはいなかったため、多様な改革の要求が無秩序に大学に押しつけられることになった（Kristensen 2010）。日本でも評価関連の諸概念の定義が混乱したまま議論が進んでしまっていることを、羽田貴史がくり返し指摘している（羽田 2004、羽田 2005、羽田 2007、羽田 2009）。

いずれにせよ、評価の制度化や実施に関しては、大学内部からの不満と外部からの不満とが、すれ違いながら渦巻いている。それゆえ、教育研究の評価をどう考えるのかについての整理が必要である。また、その整理を通して、さまざまな評価の手法や仕組みの意義と限界を確認することが必要である。評価がムダだとか全部やめろというわけではない。評価の提案はいずれもそれぞれ現実に対する部分的な合理性をもっている。何かを改善することに有益である側面はもっているであろう。だが、だからといってそれを採用すればよいわけではなく、その評価の手法や仕組みを採用することの難点や、それによって生じるかもしれない副次的なマイナスの効果などと一緒に考える必要がある。

というのも、たとえば、評価の仕組みは多ければ多いほどよいわけではないからである。日本の大学教員の生活時間や意識に関する 1992 年、2007 年の 2 時点比較調査から、福留東土（2010）は、憂慮すべき事態を描いている。1 週間平均時間（学期中）をみると、研究大学での教育時間が増加して、それ以外の大学と均質化していつているのだが、研究時間はどちらのタイプの大学でも大幅に減少している（研究大学 24.5→17.0 時間、それ以外の大学 20.7→16.5 時間）。教育に費やされる時間は少しだけ増加しているものの、大幅に増えているのは、研究でも教育でもない時間である。南部広孝（2010）は同じ調査のデータから、一般国立大学で、教員の環境への不満が急増していることを明らかにしている。この結果からすぐに「研究時間の減少が評価の水ぶくれのせいだ」といいたいわけではない。この結果から読みとるべき大事な点は、「大学教員に使える時間は有限だ」ということである。

評価は、大学の教育研究にとって、あくまでも付随的なものである。それが大学教員が取り組むべき本来の業務——教育研究——を圧迫し、機能低下に至らせるとすると、それは本末転倒というしかない。何重にも評価の仕組みが作られ、それぞれが膨大な手間やコストを要するものになるならば、それは大学にとってきわめて憂慮すべき事態である<sup>1)</sup>。

評価の仕方についての試行錯誤を続ける英国政府は、2003年に規則改良プロジェクト・チーム（Better Regulation Task Force: BRTF）を設置し、評価が煩雑すぎないようにする勧告を出す権限を与えた。「このグループは、イギリス保健サービスと高等教育を含む公共部門の組織に対して質保証と評価のプロセスを調査し、可能であればそうした組織への規則上の負担を軽くするという権限（remit）を与えられ」、そのグループのメンバーと高等教育規則審査グループ（Higher Education Regulation Review Group: HERRG）のメンバーとが、「高等教育機関にかかる負担の監視を続け、あまりにも多くの措置が導入されていると考えられる場合には、高等教育質保証機構（QAA）などの組織に対して勧告を行う権限を有している」（Clarke 2007, pp.21-22）というのである。組織論からみると「屋上屋」にもみえるが、あまりに評価があまりに肥大化・些末化して、高等教育機関の負担が増大してしまう傾向に対して、英国政府が取らざるを得なかった対応であると考え、評価を細かく、しかも何重にもやっていくことのもつ弊害は容易に理解できるだろう。

時間とお金を大量に浪費させ、かえって大学の機能低下をもたらすような「評価の氾濫」にならないために、われわれは評価という営みが持つ両義性をきちんと意識しなければならない。評価は教育研究を改善しもあるが、圧迫したり歪めたりもするのである。教育研究の評価をどう考えるのかについての整理においては、さまざまな評価の手法や仕組みの意義と限界とを確認し、適切な方法を適切な範囲で取捨選択することができるようにすることが必要だといえる。

2001年にいわゆる「遠山プラン」を出して大胆な大学改革を促進した遠山敦子・元文部科学大臣も、自著の中で大学評価が持つコストやリスクに注意をうながしていた（遠山2004）。第一に、「評価には必ずコストがともなう」。評価の負担は直接の教育研究とトレードオフの関係にあるという側面を持つ。「評価にかかるコストを上げれば、その分教育研究に直接かけるコストが減るという関係にある、ということに注意が必要である。評価の「限界効用」は必ず逡減する、ということも、まず間違いないであろう」。第二に、「評価にどんなにコストをかけても、評価には必ずリスクがともなう」ということである。「我々は、いくら精度を高めても、評価が不完全でしかあり得ないことを十分承知しているはずであるにもかかわらず、その結果が出るとそれが万能であるかのように錯覚しやすい」。それらの問題を指摘した後、遠山は、「大学評価に過大な期待を抱くことなく、しかし、大学評価の重要性を忘れることなく、私たちはそのシステムを磨き上げなければならない」と述べている。

「システムの磨き上げ」をどうしていくのか。さまざまな評価の議論や実践をどう分類

し、それぞれの可能性や限界をどう考えるのかについては、第1部第1章で詳述していく。ここでは、次節で現状の問題点を概括的に整理し、その後、スリムで効率的な評価システムの構築のために何を考えていかないといけないのかについていくつかの提案をしてみたい。われわれが基本的に主張したいのは、評価の目的と手段に関してそれぞれ議論上や実践上の混乱があることを指摘し、適切に選択することが必要だ、ということである。

### 第3節 評価の目的と手段をめぐる混乱

#### (1) 目的をめぐる

評価の目的をめぐる混乱がみられる現状に関して、ここでは二つのことを論じたい。第一に、複数の目的を同時に求めることから混乱が生じているのではないか、ということである。「目的の安易な多重性」と呼んでもよい。本来、ある目的に向けて合理的に編成されるはずの評価の仕組みが、現実にはしばしば別の多様な目的を含み込むことで、「何をなぜ、どう評価するのか」があいまいになったり、複数の目的の間で齟齬が生じたりしているのではないだろうか。

第1部第1章で詳述することになるが、われわれは多様な評価の仕組みを、目的と主体とで分類をした。評価の目的としては、「改善・向上」か「判別」かという区別ができる。評価の主体については、「内部評価」か「外部評価」か、という区別ができる。それら2通りの評価の分類を組み合わせて図にすると図1のようになる。①と④とを区別すること、②と③とを区別することができていれば、評価それ自体の目的に沿って、スリムで効率的な評価の編成が可能になる。

|    |       | 主体   |      |
|----|-------|------|------|
|    |       | 内部評価 | 外部評価 |
| 目的 | 改善・向上 | ②    | ①    |
|    | 判別    | ③    | ④    |

図1 多様な評価の仕組み

しかしながら、現実には、しばしば①と④とを両方目的に掲げたり、②と③とを両方目的に掲げたりして一つの評価が行われている。できるだけたくさんの目的を盛り込んでその評価制度の意義を強調したいという思惑が働くのかもしれない。たくさんの方向からの提案や要求に応えようとする政策的意図が入るのかもしれない。あるいはそもそも評価の視点が明確に定まっていないことによるのではないだろうか。

しかしながら、問題はここから生じる。改善・向上を目指した評価では、率直に状況の問題点を浮き彫りにした情報がよい情報である。うまくいっている点に関しては、大量の

情報を準備する必要はない。改善すべき点が明確化できれば、それで足りるからである。

ところが、改善・向上を目指した評価を同時に判別にも利用しようとする、やっかいな事態が生まれる。たとえば外部評価の場合を考えてみていただきたい。一つには、個人であれ機関であれ、都合の悪い情報は出したくないのが普通である。真に問題点を浮き彫りにする情報は出てきにくくなる。もう一つには、できればよく見せたいという意識が働くから、集約するために上がってくる情報は、それも「うまくいっている」ことが強調されることになる。特に、評判に直結し、資源配分に活用されることが予定された評価は、当事者による隠蔽や歪曲を誘発することになるのである。これらのことは組織内部で行われる「内部評価」でも同様である。教員個人の情報であれ、各部や学科の情報であれ、大学組織内の上位の部分で判別に利用されようとすると、真に改善すべき情報は隠されることになる。

M・トロウは、情報が隠蔽されがちな英国の大学改革の状況を説明して、「これは、占領された国の役所が占領者に提出する報告や、統制経済のもとで国有の企業や農場が中央政府に提出する報告に似ている。そういう場合には、真実を伝える習慣は廃れ、現場から上がってくる報告は、彼らが持ち出ししてくる根拠にもとづくことで、事実との関係がどんどん薄れていく。ラインを上がっていく情報が中央から下される評価や資源に影響するときには、報告は発見や真実を伝えるものではなく、真実を——特に報告する機関の問題や欠点をあらわにするような真実を——出し惜しんだ、宣伝文書 (public relations documents) になってしまうということを、われわれは知っている。しかし、アカウンタビリティは真実を語ることに依拠している。そこで、中心的な問題は、真実を伝えると罰を受けたり、達成を装うと褒美を得たりしないような、アカウンタビリティのシステムはどうやって作ればよいか、である」と論じている (Trow 1996, p.314)。

判別と改善という二つの目的の混用が重大な問題を生み出す典型的なものは、アウトカム測定をめぐる議論においてである。前節で、われわれは「個別化した利用法」と「一般的知見を引き出す情報としての利用法」とを区別しておいた。前者は、主に個人や組織の判別を目的としてなされる。「学生Aは大学の4年間の間にどのくらいジェネリックスキルが伸長したか(専門分野の知識が身についたか)」、「X大学は4年間の間に学生集団のジェネリックスキルをどの程度伸長させているか(専門分野の知識を身につけさせているか)」といったことが測られる。それに対して、「一般的知見を引き出す情報としての利用法」の場合には、「どういう属性要因や環境要因をもった者がジェネリックスキルが伸長しやすいのか(専門分野の知識を身につけやすいのか)」を、サンプル集団を対象にした調査で明らかにしようとするものである。

両者はまったく調査の考え方やあるべき姿が異なっている。

まず、調査結果のフィードバックの仕方が異なる。「個別化した利用法」の場合には、学生一人ひとり、あるいは組織の単位一つひとつが、点数などで表示されることになる。そこでは良いか悪いかは別にして、選抜や査定と直結した利用のされ方が可能になる。点数



を上げるためにもっと努力をするようになるなどの「正のフィードバック」もあるが、「点数のために勉強する（させる）」という学習動機・教育理念の転倒のような「負のフィードバック」も起きる（われわれは高校受験、大学受験でいやというほどこれを経験している）。それに対して、「一般的知見を引き出す情報としての利用法」の場合には、サンプル調査からえられた一般的知見は、どこでも誰でも利用可能である。厳密にデザインされた比較的小規模の調査によって、もしも退学しやすい学生の属性要因や環境要因が明らかになったとしたら、通常はどの大学でもそれを有効な対処法を講ずるのに役立てることができるはずである。

次に、誰（どの範囲まで）を調査するべきなのかという点がまったく異なる。後者では、ごくかぎられた良質のサンプル集団と入念に設計された調査票があれば、かなりのことを明らかにできる。だが、前者の場合には、対象は集団全体へと拡大すべきだという主張を呼び起こすことになる（学生全員、大学全体へと）。コストのかかり方がまったくちがうのだ。

最後に、測定しうること／してよいことが異なる。「個別化した利用法」の場合には、受検する当人（人、組織）が判別されてしまうのだから、回答は戦略的になされてしまう。自分が不利になってしまうような回答はあえて虚偽の回答をする、というふうな。「一般的知見を引き出す情報としての利用法」の場合には、そのようなことは起きにくい（正直に回答してもらえる）が、逆に、動機づけが乏しいので、真面目に回答してもらえるかどうかについて問題が残る。

測定してよいことの範囲も異なる。「一般的知見を引き出す情報としての利用法」の場合には、回答者は匿名化される（個人が特定されない）。それゆえ、調査分析のデザインによっては、かなりプライベートな事項まで質問することが許されるのが一般的である。ところが、「個別化した利用法」の場合には回答者が誰か特定されたままデータ処理されるので、プライベートな内容の質問を設定すること自体に大きな問題になりうる。もしも「個別化した利用法」の悉皆調査で、行政機関が個人のプライバシーに深く入り込むような情報を回答者がわかる形で集めたとしたら、それは市民の権利として許せないものと映るだろう。

要するに、「個別化した利用法」のアウトカム測定と、「一般的知見を引き出す情報としての利用法」のアウトカム測定は、厳密に区別されねばならないのである。

日本でも、米国にならって現役学生向けにジェネリックな能力を測定する「共通テスト」の試みが検討され始めている。新聞報道では、「対象は全国の大学。大学として参加するかどうか、何人の学生を受験させるかなどは、各大学の判断に任せる」とされている（『朝日新聞』2012年2月16日）。しかしながら、これは明らかに、今述べたアウトカム測定の区別がなされていない。サンプリングが適切になされないから、きっと大したことは明らかにならないし、各大学ごとの平均点が意味するものも判然としないものになるだろう。それどころか、大学によっては優秀な学生を選び出して受検させ、「ジェネリックスキルを伸長させる大学」という評判を得ようとする戦略を採用するかもしれない。さらには、「こ

れでは個々の学生の成長度がわからないから、全員を受けさせてしまえ」という大学が続出する可能性もある。結果的には、卒業時のジェネリックスキルの点数が一生履歴書に書かれてしまうような社会になってしまいかねない——要するに、目的を明確に限定しないアウトカム測定制度は、混乱を極めることになるということである。

評価の目的の混乱については、さらに別に問題も考えられる。評価が複数の目的を同時に遂行しようとする場合、目的ごとに異なる種類の情報が必要なため、結果的に膨大なデータの収集や分析の作業が必要になってしまいかねないという問題である。たとえば、日本における認証評価の近年の動向は、「改善・向上」をより強く打ち出す傾向にあるが、その場合、もしも一定の水準を満たすことに足る情報だけでなく、教育の実情を示すための多種多様な情報が求められるようになるとしたら、大学側の負担は膨大なものになる。あるいは、外部者に対して大学の教育の実情を伝えるためには、内部者に対してとは異なる種類の膨大な情報を付加する必要がある。また、評価を受ける側からみた場合、「うまくいっている」ことを大学の状況をまったく知らない人たちに示すためには何をどこまで情報として示せばよいのかがはっきりしないから、評価のために出すべき情報の範囲が不明確になり、膨大な作業を強いられることになる<sup>2)</sup>。

目的の異なる複数の評価が時間をおいて折り重なるという場合、もう一つの問題が生じている。「転用されてしまうこと」を恐れて率直さを欠いた内部評価なされがちだという問題である。改善・向上をめざした内部での真摯な評価（②）を妨げているのは、それがしばしば事後的にさまざまな事情から外部からの判別（④）の資料へと転用されてしまうことをおそれるからである。自分たちが作る内部評価の資料が、下手をすると資源配分や評判に直結してしまうかもしれないというふうに考えれば、当事者の内部評価が無難な資料で埋められてしまうのは、故なしとしない。当事者の心がけの問題ではなくて、評価のシステム全体の問題である。

第二に、今述べたこととちょうど逆の現象も起きていて、それも憂慮される。一つひとつは限定された目的を持つ評価が、全体として取捨選択なされないまま、無秩序に大学の中に入ってきているという問題である。戸澤幾子（2011）は、現在わが国の大学で、「組織として対応しなければならない評価が次のように重層的に存在する」とのべ、次のものを挙げている。

- ・ 自己点検・評価
- ・ 機関別認証評価
- ・ 国立大学法人評価
  - 毎年度の業務実績評価
  - 中期目標期間における評価
- ・ 専門職大学院の認証評価
- ・ 外部評価

- ・競争的資金プログラムの申請。中間・事後評価
- ・日本技術者教育認定機構（JABEE）による技術者教育プログラムの評価
- ・大学内での評価（教員評価、学内競争的資金の評価）

これら以外にも、特定分野の資格認定に関わる団体による評価、FD 関連での調査と評価など個別に対応を迫られるような評価はたくさんある。戸澤は「評価の重層性」と呼んでいるが、それは「異なる制度であっても教育研究という同じ営みを評価する以上、作業の重複は否めず、重複感、過重感が生じやすいことも確かである」と述べている（戸澤 2011, p.26）。

もちろん、それぞれの評価は明確な目的と、それに見あった手続きとが採用されるべきである。しかしながら、縦割りの行政の中で大学の教育研究の多様な側面に関してアドホックに個別分化した評価制度が作られたり、多様な学協会や民間機関などが個別に自分たち独自に設定した基準で評価を要求してきたり、大学の中の多様な事務管理部门がそれぞれ独自に評価の制度を作っていたりしたら、結果的に「評価の重層性」は非効率と疲弊を生んでしまい、本来の大学の研究教育機能が阻害されることになりかねない。このように考えてくると、多様な評価の氾濫はやっかいなジレンマを抱えていることがわかる。前項で述べたように、一つの評価に複数の目的を盛り込めば、制度自体は簡単になるが、目的に応じた手段の面で混乱が生じることになる。だからといって、別々の目的を持った評価を次々と付加していけば、「評価の重層性」による非効率と疲弊が生まれる。

考えてみれば、1990 年代以降の高等教育政策、特に 2000 年代に入って以降のそれは、矢継ぎ早にさまざまな取組がなされ、大学に多大な影響を与えてきている。しかしながら、時として何のためかよく見えてこないことも少なくない。大学評価に関わる諸政策もまさにそうである。いろいろな思惑や配慮のバランスを取るため、結果的に一つの制度の中にたくさんの目的が盛り込まれ、何を実現していくべきなのかがあいまいなまま評価がなされる結果になったり、統一性のない多種多様なアクターが思い思いの評価制度の提案や構築を行い、つぎはぎのように評価の仕組みが折り重なっていたりもしている。大学評価に関わる高等教育政策の次元で重要なことは、高等教育政策全体の構造をもっと明確にし、最低限共通に必要な制度の整備と、あったほうがよいもの／あってもなくてもよいもの／あるべきでないもの、の明確な線引きとを行い、「評価システム」とでもいえるべき体系的な構造を作り出すことであるように思われる。

ここで検討してきたことをふまえると、評価は、限定的で明確な目的を持つ必要がある。また、非効率や疲弊を生まない程度の種類に評価の数を抑制しておく必要がある。どうすればよいかは、次節で考えることにし、その前に、混乱のもう一つのポイントである「手段」について考察しておこう。

## （２）手段をめぐって

評価の際に採用されるべき手段に関しても、議論や実践のうえで、さまざまな混乱を生

み出している。その背景には、評価技法に関する研究の未成熟や専門家の不足があるだろうが、そのほかにも、目的の不明確さからくる恣意的な手段の選定、当事者たる大学における担当者の専門性の欠如、会議等における思いつきの提案など、さまざまな要因がありうるだろう。

ここではいくつかの混乱を論じたい。第一に、「ともかく数字で」という姿勢が、教育研究の現実をゆがめたり、改革努力を形骸化させている。外部者にとっても「わかりやすい」という長所があり、内部者にとっても数字さえカウントすれば膨大な質的資料を揃える必要がないという点で楽なのだが、そもそも、数字で表されているのは複雑多岐な現実から切り出された現実のごく一面にすぎないし、主張や判断の根拠としては、量的な根拠はたくさん、質的な根拠や論拠の中の一つにすぎない。教育という事象にはデジタル化できないものもあるし、すべきでないものもある（広田・伊藤 2011）。

にもかかわらず、量的根拠を求める議論が横行することで、さまざまな問題が生じている。しばしば現実の「教育研究の質」から遊離した数字が指標になることで、無意味な混乱が生じているのである。たとえば、数字になりやすいもの——たとえば資格取得率とか留年・卒業率など——が使われることで、数値化しにくい教育目標の設定や達成が等閑視されてしまうとか、カリキュラムや授業が歪んでしまうといったことが起きている。あるいは、特定の数値化された指標さえ満たせば問題ないということで、教育研究の本来の意味での改善の努力が欠如していく、といったことが起きている。

第二に、教育や研究が本来的に持っている目標の二重性を理解せず、目標の達成を直接評価しようとするために形骸化が生じている。教育も研究も、簡単には達成できない高い理想を「目標」として掲げて遂行される。しかしながら、その高い目標は、通常は十全に実現することはない。その目標は、教育や研究の実践を方向づけ、統一性を与えるものとして機能するものだからである。現実には、より低いレベルに実践的な個別的で言語化されない目標を設定して、柔軟に教育実践を展開したり、現実的に獲得可能な知見に向けて研究をとりまとめていく。しかしながら、評価においては、しばしばこの2種類の目標が区別されず、直接前者の達成を検証しようとする。ここで問題が起きる。これについては、もう少し説明が必要である。

「教育目標」という概念について哲学的な考察を行ったドイツの教育学者W・ブレツィンカは、この概念には2種類ものが含まれると述べている。一つは、「被教育者のための理想としての教育目標」である。「人格の一定の状態、人間の一定の特質に関わるもの」である。生徒や学生が「どうあるべきか」を言語化したものである。もう一つは、「教育者のための課題規範としての教育目標」である。それは、「教育者に対して、被教育者が教育目標の内容すなわち実現されなくてはならない性向構造をその人格内にほんとうに（完全ないしそれに近いぐらいに）実現するように、教育者は行為すべきことを規定する」。つまり、教師はいかにふるまうべきかを定めたものである（ブレツィンカ 1977=1980）。後者の教育目標に対して、教師は最善の努力を尽くすことはできるし、その規範はそれを要請する

のだが、にもかかわらず、前者の目標が十全に達成されるという見通しはない。非教育者にとって教育者は「他者」であるため、教育者の行為は（仮に最善の振る舞いをしたとしても）、教育の結果を十全に約束しないからである。広田照幸（2009）は、これを「教育の不確実性」という語で説明している。

研究も同様の不確実性を持っている。研究者は、いまある知と条件を踏まえて研究の終着点として高い理想を設定する。それは、達成可能性を示すのではなく、ねばり強い一貫した努力の方法を指示するための目標として設定される。しかしながら、現実の研究の遂行は苦難やアクシデントに満ちており、研究の過程で実際にはもっと現実的に達成可能な目標が据えられ、それが実現されていく。場合によっては、研究途上で目標は変更され、思いもかけない成果の産出に向かったりもする。研究において最も斬新で生産性が高いのは、皮肉なことに、そうした挫折を経て当初の目標とは異なる産出物を生んだような場合であることはしばしばである。

このような「目標の二重性」を考えると、「目標を設定する」——教育や研究を遂行する——成果を測定して達成度を測る、という手続きは、大きな誤謬を含んでいるといえる。最初に設定された高い目標は実現されないことが常態であり、それが当然なのである。いわば、究極の目標とそれに到る段階的目標とを同じ水準で考えるべきではない。究極の目標達成のための段階的試行には失敗や誤謬を含むことを否定的に見るよりも、革新は常に予定外、計画外のところで起こってきたという歴史を踏まえて、我々は挑戦を続けて行くべきなのである。

にもかかわらず「目標がどの程度達成されたかを評価する」というスキームが作動しているから、結局のところ、当事者はアウトカムについての数字をひねり出したり、虚飾的な作文をせざるをえなくなる——評価の形骸化である。

もちろん、高い理想としての目標は達成すべきものと考えないで理念レベルにとどめ、現実には達成可能な目標のみを設定し、その達成度を測るという評価のやり方はある。国立大学法人の中期目標の設定などがそれにあたる。しかし、教育研究の活動の実質に関わる部分では、高い理想としての目標の達成のために、当初設定されていた「教育者（研究者）のための課題規範としての教育（研究）目標」が簡単に捨てられて軌道転換することもありうるから、現実的な目標（の達成）のみを評価で重視する仕組みにも大きな問題が残ってしまう。

第三に、不適切な形式的要件の要請や、評価項目の非現実的な機械的細分化の問題がある。いずれも目的・目標に適合しない手段が求められるという事態である。

「不適切な形式的要件の要請」というのは、外部からチェックしやすい形式的要件を定め、それを履行させることで質を保証したと考えるやりかたである。これは、もちろん有効な場合もあるが、現実の多様性を無視して制度化されると、たちまち形骸化してしまう。

たとえば、「授業の質を高める」という目標が設定されたとして、それを「シラバスの記載方法を細かくさせる」という手段で実現しようとする場合などがすぐに思い浮かぶ。1

年先のまだ顔も知らない学生を相手にした各回の授業をどこまで詳細に準備できるかは、科目によってははなはだ不都合なことが起きてしまう。学生とのやりとりを経て、彼らの関心や問題意識にそった授業内容の展開をやっていく、といったスタイルの授業を封殺してしまう。取りうる道は2つである。授業をシラバス通りに形式主義的にやるか、あるいは当初設定したシラバスを形骸化したものにするか、のどちらかである。後者の場合（シラバスから外して学生の興味関心にそって実際の授業を組み立て直していったら）、授業アンケートで「この授業の教員はシラバス通りに授業をしなかった」という数字が出て、「ダメ教員」の烙印を押されることを覚悟しておかないといけないけれども。

「評価項目の非現実的な機械的細分化」は、文字通り、上位の一般的な評価項目が、どれかの思いつきのようなくさんの細分化された評価項目へと細分化され、教育研究のあらゆる場面をしばるような事態である。仮に1つの項目に4つの観点を設定し、それぞれに4つの指標を設定するとすると、 $4 \times 4$ で16個の評価指標が作り出されることになる。それぞれの評価指標に対して、4種類の情報を集めて説明しようとする、 $16 \times 4$ で64種類もの情報が必要になる。上位の組織から下位の組織に、また、組織内の管理層から評価担当の実務者に、実務者から各教員へと評価のスキームが下りていく過程で、こうした「必要とされる情報の項目数の水膨れ」が起きてしまっている。抽象度の高い上位の評価項目に対して具体的な事象の数や回答の数の積み上げで対応しようとするから、末端に指示が出るときには雪だるまのようにふくれあがっているわけである。

つけ加えておけば、細分化された下位項目が持つ問題はもう2つある。一つは、項目間のウェートをどう考えるのかという点で、いつも恣意性がつきまとう。この点は、国立大学法人の評価を論じた金子元久（2007, p.11）が指摘している。「上位の評価の根拠を明確にするために、これをいくつかの観点や項目にわけ、その評価を総合して上位の評価を得るという順序を踏むことになっている。しかし、その際に、下位の観点、項目がどのようなウェートで加えられるかが明確でない」と。多様な種類の大学評価で同様の問題がみられる。

もう一つは、上位の項目と下位の項目との間にズレがありうる、ということである。Aという上位の概念を構成する下位の要素として、a1、a2、a3、a4、と適切に下位項目が設定されれば問題はないが、a1、a2、a3、b4、というふうに、異質なものが混じっていることがある。某大学の理事から、「教員の授業の質を評価するのに教授会への出席率を使うのはどうでしょうかね」と尋ねられたことがある。その理事にいわせると、「教授会にまじめに出席している教員は、授業もまじめにやっているはずだ」というのである。これは明らかに下位の指標として不適切だが、おそらくそれに類するものが全国で散見されるにちがいない。

評価の手段に関わる混乱の第四は、副作用への配慮が十分なされないことからきているという点である。新しい評価のやり方を提案したり、評価の充実を訴えたりする人たちが、しばしば、副作用の問題を軽視したり無視したりしていることに、大きな問題がある。そ

の最も重要なものの一つは、時間の希少性や人材の希少性を忘れている、ということである。評価の精度自体は、評価の仕組みが緻密であればあるほどよいようにもみえるが、それは、前の方で遠山敦子（2004）の危惧を紹介したように、大学が果たすべき多様な機能の縮減を生みかねない。評価が時間コストや金銭コストに見合わない事態を心配しないといけないのである。

近年は、米国や英国などを見習って日本でも分野別評価をやれ、という議論がある（舘 2005、舘 2008、木村 2011）。しかし、もしも全大学の全分野に関してそれをやるとすると、膨大な時間と人員の覚悟が必要である。90年代の英国の経験をふりかえった村田直樹（2004）は、「分野別教育評価については、膨大な作業のわりに評価結果が全般的に高く、結果を活用した施策を講じにくいことなど、評価を受ける高等教育機関と評価結果を利用する関係者双方にとって満足できるものではなかった」と述べている。また、Clarke（2007）は、「機関、学科レベルの審査はいずれも準備に時間・労力という資源を要するため、1990年代に入って教育と学習の質保証に二つの並行するシステムが登場したことは、高等教育機関の側からは嘆かわしいこととみなされた」（Clarke 2007, pp.9-10）し、特に、「学科別審査には教師陣がこれ〔大学監査〕よりもはるかに多くの時間を傾注せねばならず、そのために関係者すべてに許容される度合いはかなり低かった」（同論文, p.10）と述べている。

また、英国の大学で実施されてきている学外試験を理想視し、わが国にも導入すべきだという議論があるが、Clarke は、「少数派に属する分野や教師陣が高等教育以外で高い収益力がある場合など、学外試験委員を採用することが難しくなっている一部の学科もある。こうした理由のほかにも、教師陣は研究、教育、行政事務という彼らの仕事のあらゆる面で圧力に晒されており、そのために追加の献身を引き受けるには気が進まないことがあるだろう」（同論文, p.14）と述べている。優れた評価者は、多くの場合優れた研究者や教育者でもあり、評価制度の「充実」は、評価というメタ的な活動のために、貴重な人材による教育研究の活動を妨げる側面があることが意識される必要がある。

#### 第4節 選択と自律性の必要

現在の評価の仕組みが最善とはいえないだろう。だからといって、ただひたすらもっと緻密化したり、新しいものをつけ加えていけばよくなるというわけでもなかろう。かえって、評価システムの効率性やコスト・パフォーマンスが低下したり、大学が果たすべき本来の機能にしわ寄せが生じるような事態にもなりかねない。ここまでの検討を踏まえて、われわれが必要だと考えるのは、スリムで効率的な評価システムである。その構築のために何を考えていかないといけないのか。いくつかのことを提案したいと思う。

### (1) 「NPMの失敗」への配慮を

大学評価の制度の導入や拡大を支えてきたのは、行政学で NPM (New Public Management) と呼ばれる思想である。それは民間の経営手法を大学を含めた公的部門に持ち込むことによって効率性を高めようとするもので、南島和久 (2010) の整理によれば 2つのタイプがある。一つは、「規制緩和、民営化といった政府の役割の見直し、あるいは PFI (Private Finance Initiative)、規制改革、民間委託、あるいは指定管理者制度、市場化テストなどの「市場メカニズムの活用」を主張する」ような「市場型 NPM」である。もう一つは、「民間企業の目標管理制度の導入、成果主義や人事評価、業績給、公会計の改革や内部統制・外部監査などの諸改革」のような「企業型 NPM」である (南島 2010, pp.25-26)。国立大学の法人化は前者の文脈に、大学マネジメントに PDCA サイクルを導入するような動きは後者の文脈に位置づけられよう。

「市場型 NPM」においては、NPM 改革と並行して「評価官僚制」が形成され、評価の拡散現象、すなわち監督・監視が強化される「統制の多元化」が進む、と南島はいう。また、「企業型 NPM」では、より直接に評価が「中心に位置づけられ」、目標管理や人事管理に活用される (同論文, pp.26-28)。

大事なことは、「国家の失敗」「市場の失敗」と同様に、「NPM の失敗」も起きうることである。増殖した評価官僚制がはらむ官僚制的硬直化、戦略的な目標設定による目標管理のゲーム化、多元化した統制の錯綜による非効率の発生などを思い浮かべてみればよいだろう。南島は「企業経営では経常利益の増減が最終的な規律点となる。しかし、公的部門はどのポイントに基づいて判断をするべきなのかという点が曖昧である。しばしばその価値は複数でありこれは相互に錯綜する。そして、究極においてこれを規律する単一のドクトリンが存在しているわけではない。この点こそが公的部門の管理の難しさであり、この点に対する配慮が NPM 論者の重要な欠落点になっているのである」と述べている (同論文, p.23)。目的の多重性に由来する混乱や評価の氾濫などは、このような NPM に固有の弱点に由来するものであると考えられるから、NPM の手法の洗練や徹底 (さらなる評価制度の新設や拡大) によって解決するようなものではない。むしろ、現実には照らして絶えずその手法の氾濫に限定をかけていくことが必要である。ともかく、マネジメントやガバナンスに関して評価に大きなウェイトを置く NPM の仕組みがシステムを効率化する唯一絶対の「解」ではないことを、前提として考えておく必要があるだろう。

大場淳 (2009) の整理によれば、確かに NPM の実験は、いくつもの失敗を経験してきている。思いきった市場化や成果主義の資源配分を行ったチリやニュージーランドの高等教育改革は失敗だったというふうにいわれている。評価の仕組みが大学に浸透するにつれ、無難な模倣ばかりが流行して多様な教育研究が衰退してきている。そもそも、大学の活動の業績を数値化して示すことの困難性がクリアされてきていない。また、「自律性拡大の利益を享受するのは主として大学執行部で他の多くの教員はむしろ自由を失い、自由を拡大できる教員は企業的活動を行う一部の者に限られる」 (大場 2009, p.191)。さらに、大場



は、OECD が 2008 年から、NPM 路線の見直しに向かっていることを紹介している。

1990 年代以降 NPM の行政への適用を加盟国に強く推奨してきた OECD は、最近の報告書 (OECD 2008) で、研究資金の配分方式が競争的資金に重点を置くに伴って様々な問題が生じているとし、短期的な成果を挙げることに寄与するものの長期的観点からは研究確信の発展が阻害され、リスクの高い計画が取り組まれない結果研究活動の同質化が進み、斬新な発見がなされる可能性が低くなると指摘している。加えて同報告書は、資金獲得のために膨大な時間と労力が費やされ、後継者育成が阻害されていることにも懸念を示している」(同論文, p.191)

評価システムや評価手法の細部を議論する前に、まずは、近年の評価の制度化の前提になってきていた NPM の手放しの礼賛に対して、少し冷静になってみる必要がある。おそらく、南島のいうように、国家の介入や市場原理と同様に、NPM もまた、「どの程度までその原理で制度を作るのか」という選択が働かないといけないということである。

## (2) 収斂しないシステム

もう一つ必要なのは、多元性が保証されるシステムであることが必要だということである。

国単位でいえば、外国で取り入れられている仕組みを理想視して無批判に追随していくのではなく、日本の大学の歴史や現状を踏まえながら、適切だと考えられるものを適切な形で導入し、日本の法体系や組織風土にあった形でリメイクし、定着させていく努力が必要であろう。国内レベルでいえば、それぞれの大学が置かれた状況や、それぞれの大学がめざす方向によって、教育研究の評価には多様な基準が存在していることが必要だろう。

この点に関しては、ヨーロッパの動きが注目される。高等教育品質保証期間国際ネットワーク (International Network of Quality Assurance Agencies in Higher Education: INQAAHE) の理事を務めた (その後会長) Richard Lewis (2005a) は、「1997 年 (リスボン協定 Lisbon Convention—筆者) までの議論の焦点は同等性にあり、同等性は厳密でなければなりません。ある国の学位をもっている場合に、その学位は別の国の学位と正確に同じでなければなりません。現在では、比較可能性 (comparability) という、より微妙な概念に移ってきています。どうすれば比較できるのか、何が共通しているのか、何が違うのかに着目するわけです。1997 年とは根底が異なっています」(Lewis 2005a, p.87) と述べている。高等教育の資格の枠組み (qualification framework) も、完全に内容が同じであることを求めているわけではなく、「ある特定の資格を有するとはどのような行為であるのか、学位を取得したことは何を意味するのか」について、「それぞれの国で同じとは言えないけれども、理解できる仕方で表現された資格の枠組みを作しましょうということで」(同論文, p.86) ある。ボローニャ・プロセスにおける質保証の枠

組について、大場淳（2007）は、質保証のポイントは相互の信頼であって、拘束性の強い共通枠組みではないことを強調している。ENQA が 2005 年に作成した高等教育の質保証に関するガイドラインの影響について、ウェスターハイデンは、ヨーロッパじゅうの大学教育が統一化されてしまうようなものではないと説明している。「ガイドラインにおける standards は教育の中身には関わらない。それは、……すべての機関に質保証の仕組みを持つよう求めるものにすぎない。内部質保証が重視され、それによってカバーされる領域の最低限の定義がなされているにすぎない。各国、各機関は、多様な水準、多様なやり方で質保証をしていくことになる」（Westerheijden 2007, p.89）。

つまり、1997 年までは各国の高等教育が厳密に同等であることが追求されていたけれど、それ以降は、比較可能性（comparability）へと考え方が転換しているのである。国際的に同じ方法での質の保証が必要なわけではなく、それぞれの国で信頼に足るやり方で質の保証がなされていればよいということを意味している。実際、2003 年に INQAAHE に加盟する 60 機関を質問紙調査したものによれば（Lewis 2005b）、現実に採用されている質保証の仕組みもやり方も多様だということがわかる。所管でいうと、政府 26%、独立・政府主導 29%、高等教育機関 21%、それらの混合型 12%、専門団体 12%、である。質保証の対象は、機関 17%、プログラム 37%、両方 46%、結果の公表も、「広く公表」が 41%、「公表するが広く配布せず」が 24%、「非公開」が 35%となっている。

「ある国において、全部または一部が直接的に他国に由来する教育が提供される」場合、すなわち、「国境を越えた教育」においても、「政府は、自国の状況と要求に最もふさわしいモデルを採用する必要がある」とされている（ユネスコ・APQN 2008, p.73）。それぞれの国の状況に応じた規制の仕方を採用すればよいわけである。ここでもまた、評価の仕組みや規準がグローバルに一元化するわけではないことになる。

このようにみてくると、どの国でも、それぞれに固有の事情を踏まえた多様なオプションを採用することができる。日本の文脈をふまえた制度を作れば（そして海外に説明して信頼されれば）よいということである。大学の教育研究の評価のあり方は、この点を見失わないで議論されることが必要であろう。

国内に目を転じてみると、ここでもまた個別の大学や分野が置かれた文脈によって、教育研究は多元的に存在しうるし、またそれが必要でもある。この点は、中教審でも十分理解されている。「機能別分化」という語がそれである。「我が国の高等教育の将来像」答申（2005 年 1 月 28 日）では次のように 7 つの機能が提示されている。

高等教育機関のうち、大学は、全体として

- ①世界的研究・教育拠点
- ②高度専門職業人養成
- ③幅広い職業人養成
- ④総合的教養教育

⑤特定の専門的分野(芸術, 体育等)の教育・研究

⑥地域の生涯学習機会の拠点

⑦社会貢献機能(地域貢献, 産学官連携, 国際交流等)

等の各種の機能を併有する。各々の大学は、自らの選択に基づき、これらの機能のすべてではなく一部分のみを保有するのが通例であり、複数の機能を併有する場合も比重の置き方は異なるし、時宜に応じて可変的でもある。その比重の置き方がすなわち各大学の個性・特色の表れとなる。各大学は、固定的な「種別化」ではなく、保有する幾つかの機能の間の比重の置き方の違い(大学の選択に基づく個性・特色の表れ)に基づいて、緩やかに機能別に分化していくものと考えられる。

そのうえでこの答申では、「こうした大学全体としての多様性の中で、個々の大学が限られた資源を集中的・効果的に投入することにより、各大学の個性・特色の明確化が図られるべきである」とされている。どういう方法でどこまで政策的に個性化・特色化を進めべきなのかについては、議論のあるところではあろう。

この際に考えるべきことは、大学がいくつかの分類に本当に分けうるのであろうか。「大学」という以上は共通の基盤があるのであり、それを如何に具体的に事業化としていくか、というところで大学の多様性、もしくは機能分化があると考えべきではないだろうか。答申でも明示しているように七つの機能は例示であって本質的なことではない。社会貢献なき国際的拠点はあり得ない、というのは、むしろ先端的研究を行っている国際的拠点の現場の意識であろう。そこで、大学の大学たる基盤は何か、といえ、(1)学術の継承であり、(2)学術の創出である。いわば、従来の教育と研究の活動に対応する。しかしながら、21世紀に入り、大学のユニバーサル化と政治、経済、産業のグローバル化によって、大学における教育と研究を社会との関わりで考えデザインすることが求められてきている。これは(3)学術の「社会化」、もしくは学術の「公共性」というべきもので、大学の第三の基盤として考えていかなければならない。敢えて言えば、中公審が例示する七つの機能とは、これら三つの大学共通の基盤をどのように具体化するかの段階で様々な可能性があることを例示したことになる。それゆえに、大学の個性、資源、ミッションに応じて、上記三つの大学の基盤が機能するように、多様な活動を自らデザインすることが求められるのである。

グローバルな高等教育システムの連携や交流を考えると、比較可能性(comparability)を確保できる評価のあり方が最低限の条件となる。だが、同時に、日本の固有の文脈や各大学の固有の文脈に即した評価がなされるよう、国際的にも国内的にも過度な標準化や一律化を避ける必要がある。包括的にシステム全体に求める手段は最低限にとどめ、健全な改善工夫の自主努力と長い目で見たときの質改善の競争とがうまく作動するように考えねばならない。われわれは、多様な大学間の差異を「生かすべき多元性」とみなし、狭い枠組みでの評価がそれを一元化したり序列化したりしないような工夫が必要である。

### (3) 評価の目的の限定、限定された目的に適合した手段の選択

第3節第1項で論じたように、評価の目的に関して問題が生じていることが、混乱の大きな原因となってきた。1つの評価にたくさんの目的が盛り込まれ、適切な範囲で手段を設定できない。もう一方で、多種類の評価が調整されないまま多重に制度化されてきてしまい負担感や非効率を生んでいるといったことである。これをどう考えればよいか。

前者に関しては、評価の目的をできるだけ限定し、別の目的のためには別の評価をという切り分け（分節化）が必要であろう。どうしても2つ以上の目的の切り分けができない場合には、どれかを主とし他の目的は付随的・副次的なものとするという、目的間の序列づけを明確にして、それら付随的・副次的な目的に関しては厳格な達成を求めない、というふうな工夫が必要である。

館昭は、認証評価機関による評価の目的のあいまいさとその問題点を指摘して、次のように述べている。

ここで問題なのは、適格認定でなくてもいいということが、適格認定だけではいけないということにつながる点である。……

もし答申の線にそった適格認定を主とし、それが結果として大学の自己改善につながるという制度であれば、基準を満たしているかどうかに関心を絞った評価ができる。受ける側も、自己評価を、基準を満たしているということの証明に集中できる。結果の公表は適否にかぎり、適否に直接関係しない評価を通じて明らかになった諸点は、大学にのみ伝えるということですますことができる。／それが、自己改善と社会的な評価の獲得のためとなると、評価の範囲は一挙に拡大するのである。また、目的が違えば、本来方法も異なる。それを、複合的な目的で行うのだから、評価は複雑となる。評価結果も、適不適以外に、評価のすべてを結果として社会に公表する必要が起ってくる。／こうした複雑な評価に要する、人的、物的、時間的なコストは、膨大である。……

こうした中で、認証評価をやり抜くためには、その目的を絞って実施していく必要がある。そして、その絞り込む先は、適格認定でなければならない（館 2008, pp.7-8）

認証評価の目的を適格認定に限定すべきだという館の主張に対しては別の意見があるかもしれないが、複合的な目的の設定のために混乱やコストの問題が生じているという館の指摘には異論はあるまい。もしも各大学が教育研究の細部にわたって資料を準備し、その具体的で細かな部分まで認証評価機関が改善向上に向けた意見を出すことを重視するならば、評価される側も評価する側も大変な作業を強いられることになる。特に、必要な情報の範囲が確定しないから、個別の学科等から資料を積み上げていくのは途方もない手間がかかる。そうであるとすると、所定の基準を満たしているのかどうかを確認することを主とし、それ以外は付随的なものと位

置く考え方が適切であろう。自己改善の支援という意味では、各大学で改善の努力がどのようになされているのかについての概括的な資料をもとに認証評価機関がそれを確認するというやりかたにとどめるのが、現実的だといえるだろう。

無秩序に多種類の評価がおり重なってしまうという後者の問題について、どう考えればよいのだろうか。重複する評価の種類を整理したり、学協会や民間のプログラム評価など乱立しやすいものを抑制するやり方や、定期的な評価の間隔を充分にとって負担を減らす（頻度を減らす）といったやり方もある。特に、大学に対する外部からの評価の数を増やさない工夫が求められる。

しかし、何よりもそれぞれの大学の側に選択可能性〔任意性〕をもっと発揮できるような仕組みを工夫し、「選択する」ということを活発化させることが必要なのではないだろうか。たとえば、最低限必要な共通の評価（国際的に「学士」の質を説明するものなど）以外はできるだけ「判別」のための評価をスリムにし、どういう改善・向上に向けてどういう評価をするのか／受けるのかについて、大学や部局の選択が尊重されるような仕組みを長期的に模索していく必要があるだろう。

これまでの大学改革の中では、多くの大学は外からやってくる評価を受け身で対応するので精一杯であった。一部の大学を除くと、多くの大学は、教育研究のあり方に関する長期的な戦略や明確な社会的使命を意識することなく、評価に対応してきた。その結果、評価と名が付くものにはすべて対応する、ただし、それに関わるのは一部の担当教職員のみで、教育研究の改善・向上にはつながっていかない、という状況がしばしば起きてきた。

今後必要なのは、自分たちの組織が教育研究に対してどういう方向性をもつのか——「ミッション」と呼んでおく——を明確にし、それに向けて「大学」という自律的な組織を再構築していくことではないだろうか。大学の自律性が強化され、評価の目的、評価結果の使い方を自ら明確にした上で、必要な技術、情報収集力を身に着けることなど、大学の自己評価力が高まっていけば、これまでの評価に見られたような形式主義ではなく、外部の評価者にも社会の人たちに対しても、責任と自信を持って自分たちの教育研究の意義や重要性を説明できるようになる。

この点をもう少し考えるために、ステンセイカーとハーヴェイによる「アカウンタビリティ (Accountability)」という語のとらえ直しの提起を紹介しよう (Stensaker and Harvey 2011)。ステンセイカーらは、アカウンタビリティの仕組みには2つの考え方がある、という。一つは、「何かに確実な根拠や理由を与える」というものである。これが一般的にイメージされやすいアカウンタビリティの仕組みである。しかし、彼らはもう一つの別の考え方を示す。それは、「民主社会における継続的な対話の一部分と考える」というものである。彼らはボーベンス (Bovens 2007) の議論に依拠しながら、対話関係が構造化されている次の状況のとき、アカウンタビリティが満たされているとしている。すなわち、a) アクターと公共の場 (forum) との間に関係があり、b) その関係の中でアクターは義務を負う。c) その義務とは説明や正当化をすることであり、d) 説明や正当化されるのはアク

ターの振るまい (conduct) である。e) 公共の場は質問をし、f) 判断を下すことができ、g) アクターは、その結論 (consequences) を手にする、という7つである (Stensaker and Harvey 2011, p.15)。

「アカウンタビリティ」を、公共的な対話と考えるこの視点は重要で、興味深い。大学が自律性を確立して、評価を使いこなす能力を高めるということは、自らが提供している教育研究の質に関して継続的で公共的な対話のツールとして、評価を活用することであるということができるからである。評価の結果に一喜一憂するのではなく、継続的に社会に対して責任を持って教育研究を見直していく契機として、評価が位置づけられることになる。

このような視点から、(4) と (5) で、今後の評価のあるべき方向について明確にしていきたい。

#### (4) 内部質保証の重要性

高等教育の質保証をめぐる近年の議論は多様なアクターによって、ますます拡散しているようにみえるけれども、総じていうと、内部質保証の重視の方向に向かっているように思われる。

たとえば、先に触れた ENQA のガイドライン (2005 年) は、内部質保証を重視したものである (大場 2005、大場 2011、Westerheijden 2007, Kristensen 2010)。「ENQA の規準は、……各国の独自性を尊重する一方で、各国内では高等教育機関が自律的に評価に取り組むことを重視し、自己評価が最も重要であることも示した」(大場 2005, p.60)。英国でも、重要なコンセプトは「質保証」(Quality Assurance) から、「質向上」(Quality Enhancement) へとシフトとしてきている (Filippakou and Tapper 2008)。

日本でも、内部質保証の重視に動きつつある。大学基準協会は、認証評価の第2クールに向けて大学基準協会の新対応として、内部質保証を重視する姿勢を明らかにしている (工藤 2010、納谷 2011)。「当協会の設置目的 (大学の自律・自主の支援) に照らし、認証評価制度は「最低限の質保証のための制度」と位置づけるのではなく、「質の向上をめざすための制度」として機能する側面にベクトルを置くことを明確にし、そのための制度設計・運用を見直した」(納谷 2011, p.25) というのである。

こうした流れについて、金性希ほか (2009) は興味深い発展段階図式を提示している。金らは認証評価を各教育機関がどの程度内部の改善に結び付けていっているのかを実証的に検討したうえで、3段階の「評価の内部浸透段階仮説」を示している。それによれば、まず「自己評価を通じて教育・研究を改善する体制が存在せず、組織として機能していない」段階がある。次いで、「内部での常設の評価体制が設置される」段階があり、最後に「評価が日常的な改善体制と一体化する」段階があるというのである。金らは、「対象校が評価経験を積み、第2段階、さらには第3段階へと進むことで、対象校内部で日常的に課題を把握し修正することが行われるようになれば、前述のように新たな視点が得られる

ことは減り、第三者評価の費用対効果は逡巡せざるをえない。その場合には評価システム自体を変更していくことが求められよう」（金性希ほか 2009, p.37）と述べている。こうした観点に立てば、1991 年以降の評価の制度化の進行が一定段階に達することで、内部質保証の方向に重点が移っていくことをうまく説明できるように思われる。

もちろん、内部質保証は、自らを不断に点検・評価するという意味で、重い責任と負担を、個々の大学や教員に課すものである。米国の質保証の歴史的変遷をたどったイーウェルは、1990 年代にパフォーマンス測定（Performance Measure）が流行した時代があったが、a) まだこの時期はパフォーマンス尺度が一般的なものとどまっていたし、b) 統計上の数値を上げるデータの見せ方の工夫の余地があった、c) 基準を割り込んでも資金配分への影響は小さかったため、個々の機関の負担は小さかったという。そして、2000 年代以降の「質重視の時代」（1980 年代に続いて 2 回目）になると、特に学生の学習の質に関する評価関連で、個々の機関の負担が増加し、膨大なコストを支払うことを余儀なくされていることを指摘している（Ewell 2007）。

また、教育の改善・向上を主要な目的とした内部質保証を認証評価など外部からの評価に接続して制度を構築するとすると、認証評価機関の間での基準の不一致、設置基準と評価機関独自の基準との関係などをどう考えるかという問題も浮上する（前田 2009）。

さらに、「何が改善・向上なのか」をめぐって組織内部の成員間での考え方のちがいを顕在化させ、葛藤を生み出す可能性もある。前述したとおり、「教育研究のあるべき姿」は、分野によっても考え方が異なるであろうし、同じ分野の教員間でも大きく異なっているはずだからである。

そうしたいくつかの問題や困難はあるけれども、本稿で整理してきたような評価に関わるさまざまな問題は、内部質保証の重視によって軽減されることは期待できるはずである。たとえば、「今、うちの組織で何がうまくいっていないのか。何をめざしていくべきなのか」を当事者が率直に議論することを何よりも重視するならば、外形的で形式的な尺度を用いた評価や機械的に細分化されたたくさんの項目による評価などではなく、もっと実質的な意味をもつ評価をめぐる議論が組織の中でなされることになるだろう。

外部に向けて発信すべき情報をむやみやたらに集めて整理するのではなく、本当に必要な情報を集めて議論がなされることになるだろう。アウトカムの測定も、内部での検証に必要な範囲で利用しうる有用な道具の一つに限定されて位置づけられることになる。内部質保証を重視した評価はまた、評価の目的を明確にし、手段を明確に規定することになる。自分たちが議論するための材料を集めるわけだから、ムダなことは極力省略されるはずだからである。

このように考えれば、現場の第一線教員からは重荷 (burden)、官僚制 (bureaucracy)、ゲームプレイング (game playing) として経験されているような評価関連の仕事は、実質的な大学教育の改善向上という正当な目的と結び付けられ、もっとポジティブな意味をもつものとなるであろう。

M・トロウは 1996 年の論文で、外部に対するアカウンタビリティを重視した改革に強い批判を浴びせ、大学内部での当事者による改善・向上こそが重要である、と主張している。「教育によるアウトカムを測ることができないからといって、教育機関で何があまくやれていて何があまくやれていないのかについて、われわれが学ぶことができないわけではない。それこそが、内部でのレビューを通した内部でのアカウンタビリティの活動である。もし、内部のレビューと評価が、外部の適格判定者のそれよりももっと妥当で実りあるものだとするれば、高等教育機関は自分自身やその部局に対し真剣で周期的なレビューを行うことが必要であり、そのことが本当の効果や成果を得ることになる」(Trow 1996, p.322)と。

研究大学で研究をしない教員の問題についても、「これは外部者の audit や内部のレビューによってモニターされ、学部や大学で問題にされるべき問題である。中央政府の財政措置で有効に対処できる問題ではない」(同論文, p.317)とあっさり片づけている。そして、外部からのレビューは、その内部での努力に対して向けられるべきことを主張している。「真剣でしっかりとした内部の質のレビューは、外部の audit——教授の質やアウトカムについての audit ではなく、自己点検 (self-study) や自己批判の当を得た手続きとそのような内部のレビューが実際に与えている効果についての audit——を通して、モニターされうる。これが内部のレビューと外部のレビューとをリンクさせ、相互に支え合うやり方である」(同論文, p.322)——まさに近年の質保証重視へのシフトへの動向を先取りしていたといえる。過剰で意義の薄い評価の仕組みや統制の仕組みを整理して、本当に実質のある改善のための評価こそが今後めざされるべきであろう。内部質保証重視の動向は、そういう意味で歓迎すべきものである<sup>3)</sup>。

### (5) 「評価のシニシズム」を超える必要

教育研究の評価のよりよいあり方を考えようとして、ここまで4つの点を挙げてきた。それらに加えてもう1つ挙げたいのは、大学の教育研究のあるべき方向が評価のあり方によって決められてしまうというような転倒した事態——これを「評価のシニシズム」と名づけよう——にならないようにしていく必要があるということである。この「評価のシニシズム」という概念は、近年の事態を表現するためにここで導入された概念なので、もう少し説明しておかねばならない。

教育哲学者の田代尚弘や小玉重夫は、教育目的や理念の実現が追求されるのではなく、手段・方法の効率化や有効性のみが追求されるような事態を、「教育のシニシズム」という語で表現している(小笠原編 2003、小玉 2003)。簡単にいえば、『教育の目的』とは切り離されたところで、目先の現実的効用の追求や達成のみが、教師たちの目標になってしまう」ということである(広田 2009, p.108)。手段がそれ自体目的化してしまっている点で、それは本末転倒である。

大学における教育研究の評価をめぐる取り組みも、同様の事態を生む危険性がある



る。本来はまず先に「教育研究が目指すもの」あるいは「それぞれの大学の目指すもの」があり、それがどの程度実現されているのか、どこに問題があってまだ不十分なのかを確認するために「評価」という活動があるはずである。しかしながら、しばしば起きるのは、評価で高い評価を受けることが日常の活動の「目的」になり、評価に盛り込まれた項目で高い点数をとれるように活動の焦点が設定されていく、といった事態である。それは、評価の仕組みや仕掛けが教育研究や大学の目指すものを決めてしまう、という意味で本末転倒した状態である。目的が評価のあり方を決めるのではなく、評価のあり方が教育研究や大学の目指すもの（目的）を決める、というふうになるからである。「評価のシニシズム」は、ひたすら評価の際の項目への対応に向けて教育研究を組織しようとする大学を生み出すことになる。

考えてみれば、大学評価で論じられている目的——「改善」にせよ「判別」にせよ——、それは実質的な教育研究の中身を具体的に方向づけるものではない。たとえば、「改善」という語は、何を指して「改善」というのか、それを「改善」と呼べるのはなぜなのか、と考えると、中身を充填していない空箱のような概念であることがわかるだろう。そこに、「効率化」とか、「社会に対するアカウンタビリティ」とか、「ステイクホルダーへの情報提供」などといった中身を入れてみたとしても、今度はそれら自体が、あるべき教育研究の姿や方向やその正当性に関して何も言っていないことがわかるはずである。「質保証」といわれるときの「質」もまた、空箱のような概念である。「就職率を上げる」とか、「国際ランキングの順位を上げる」などというのも同様である。それによって、その大学がどういう意味でわれわれの社会に対して有益な貢献をすることになるのか、その大学の教育や研究がどういう公共的な使命に向けてなされるのか、そういった点の説明がまったく欠落している。

このような「評価のシニシズム」を回避するためには、評価の目的とは別の次元で、大学が果たすべき使命や教育研究が目指すべき方向——大学（の教育研究）の目的——が、個々の大学や個々の学問の論理などから内在的に設定される必要がある。では、それはどのようなものか。

狭い特定の社会的ニーズに答えるような卑近な目標を設定することはできる。「資格を取得した学生をたくさん送り出す」とか、「〇〇学の国際的な研究拠点をつくる」とか。しかし、それでは「評価のシニシズム」は避けられても「教育のシニシズム」の外には出られない。「なぜそれを目指すのか」とあらためて問いかけられたとき、答えられないからである。

大学の中から発する教育研究の内在的な目的があるとする、それは各大学や各教員が、自らの活動にどのような公共的な使命を与えるのかが問われているのではない。「公共的な使命」を広い世界に向けて説明しようとする、21 世紀の社会において果たすべき役割を、大学や大学教員自らが定義していく必要がある。そこ

では、学問が持つ普遍的でイノベーティブな性格を基盤として、その教育が果たす意味、研究が果たす役割が、社会に向けて説明されるものでなければならないだろう。

大学が外部から押しつけられるきまぐれな「ニーズ」に振り回されず、評価の仕組みに埋め込まれたアドホックな評価尺度にも従属せず、内在的に自らの公共的な使命を明確化し、それに沿って教育研究を組織していくことが必要である。「学問の自由」や「大学の自治」という古くさい語が、新たな意味を持つのはこの点においてである。大学は、自律的に自らの存在理由を作り出していく必要がある。それは、大学が自律的に自らの目的設定をすることによって初めて可能になる。自らの公共的使命を内在的に発出する大学の自律性こそ、「教育のシニシズム」や「評価のシニシズム」を超える唯一の源泉なのである<sup>4)</sup>。

## 第5節 おわりに

本章では、大学の教育研究の評価をどう考えるかについての概括的な考察をおこなってきた。ここでは簡単にそれをまとめておきたい。

まず、第1節では、日本における大学評価が不可欠になってきた社会的・歴史的文脈を検討した。グローバル化が旧来の大学に変化を迫っていること、高等教育進学者の広がりが多様な大学の実態を生み出していることなどから、大学の名にふさわしい教育とはどういうものであるかを明確にしつつ、それを実際に提供していることを社会に向かって責任をもつ必要が生じてきていることを述べた。

第2節では、評価の仕組みは多ければ多いほどよいわけではなく、効率的であることが求められると同時に、大学の多様な機能や活動を阻害しないように、スリムであることも求められるということを述べた。

第3節では、スリムで効率的な評価システムを構築するために何を考えておかねばならないのかを論じた。評価の目的と手段に関してそれぞれ議論上や実践上の混乱があることを指摘し、適切に選択することが必要だ、ということである。

目的に関しては、①複数の目的を同時に求めることから混乱が生じているのではないか、②別々の目的の評価が無秩序に持ち込まれて混乱や疲弊を生んでいる、といったことを指摘した。そこには、やっかいなジレンマが存在している。

手段に関しても、議論や実践のうえで、さまざまな混乱が生じている。①「とにかく数字で」という姿勢が、教育研究の現実をゆがめたり、改革努力を形骸化させている、②教育や研究が本来的に持っている目標の二重性を理解せず、目標の達成を直接評価しようとするために形骸化が生じている、③不適切な形式的要件の要請や、評価項目の非現実的な機械的細分化が、混乱を生んでいる、といった点を指摘した。

そのうえで、第4節では、何が考えられるべきかについて、いくつかの提言を行った。

①「NPM の失敗」に配慮すべきこと、②システムは収斂しないということを前提にして評価のあり方を考えるべきこと、③評価の目的を限定し、その限定された目的に適合した手段が選択されるべきこと、④内部質保証の方向が重要であること、⑤「評価のシニシズム」を超えるべきことを論じてきた。

## 注

- 1) 認証評価制度の導入が私立大学にどのような影響を与えているのかをインタビュー調査などによって考察した横山和子(2007)は、事務的に手間がかかるだけの「義務」であると回答する部局代表者が多いことを明らかにしている。
- 2) 認証評価における内部質保証へのあり方を考察した宮浦崇ほか(2011, p.155)は、「評価項目において、機関のコンプライアンス (compliance) と、これとは対比・対照的な質の向上 (quality improvement) を混在させることは、異なる根拠資料 (evidence) と評価指標を同時に使用した自己評価と第三者評価が行われることを意味する」と指摘している。その通りである。
- 3) 江原武一(2010)は、1990年代以降の米国の大学評価の特徴と改革動向について考察し、大学の自己点検・評価をベースにした仕組みこそが求められると主張している。その論拠は、研究評価も教育評価もまだ未開発な段階であり、それを用いるのは適切ではないという現実的な論理である。「一般的に公認されるような大学評価の原理や手法が開発されるのは、当分先のことになるだろう。それゆえ自分の大学にふさわしい大学の自己点検・評価を構築し、着実に実施していくことこそが大学の将来にとって建設的なものになる」(江原 2010, p.122)という。内部質保証の議論が尊重されるべき大きな理由といえるであろう。
- 4) 田中弥生(2010)は、企画立案と評価とを組み合わせる「評価可能性のアセスメント」を提案している。目標設定の適切化が図られるだけでなく、「何をめざすのか」についての内部での議論を誘発し、大学の自己評価力の向上させる点で注目すべき提案である。ただし、投入できる資源が制約された中では、「ある程度到達可能で現実的なレベルまで目標値を引き下げる」(田中 2010, p.36)だけに終始しかねないし、何よりも、もしも達成可能な目標のみに組織の努力が傾注されるなどの「評価のシニシズム」を生み出してしまいかねない点に留意が必要である。

## 引用・参考文献

- ブレツィンカ・W、1977=1980、『教育科学の基礎概念』(小笠原道雄ほか訳) 黎明書房。
- Bovens, Mark, 2007, “Analysing and Assessing Accountability: A Conceptual Framework,” *European Law Journal*, 13(4): 447-468.
- Clarke, Gill、2007、「イギリス高等教育における質保証」(吉川裕美子訳)『大学評価・学位研究』6、3-24。

- Ewell, Peter, 2007, “The ‘Quality Game’: External Review and Institutional Reaction over Three Decades in the United States”, Don F. Westerheijden, Bjorn Stensaker, Maria Joao Rosa, (eds.), *Quality Assurance in Higher Education: Trends in Regulation, Translation and Transformation*, Dordrecht: Springer-Verlag.
- 江原武一、2010、「アメリカにおける大学評価の改革動向」『立命館高等教育研究』10、113-124。
- 福留東土、2010、「研究と教育の変化」『IDE 現代の高等教育』519、38-41。
- Filippakou, Ourania and Tapper, Ted, 2008, “Quality Assurance and Quality Enhancement in Higher Education: Contested Territories?,” *Higher Education Quarterly*, 62(1/2): 84-100.
- 羽田貴史、2004、「大学組織の変容と質的保証に関する考察」『COE 研究シリーズ 8 高等教育システムにおけるガバナンスの組織の変容』広島大学高等教育研究開発センター、1-18。（再録：2009、塚原修一編『リーディングス日本の教育と社会 12 高等教育』日本図書センター、33-50。）
- 羽田貴史、2005、「高等教育の質保証の構造と課題—質保証の諸概念とアクレディテーション」『COE 研究シリーズ 16 高等教育の質的保証に関する国際比較研究』広島大学高等教育研究開発センター、1-13。
- 羽田貴史、2007、「行政改革における評価の動向と認証評価」『COE 研究シリーズ 28 高等教育の質的保証に関する国際比較研究』広島大学高等教育研究開発センター、1-23。
- 羽田貴史、2009、「質保証に関する状況と課題」羽田貴史・米澤彰純・杉本和弘編著『高等教育質保証の国際比較』東信堂、3-19。
- 広田照幸、2009、『ヒューマニティーズ 教育学』岩波書店。
- 広田照幸・伊藤茂樹、2011、『教育問題はなぜまちがって語られるのか？—「わかったつもり」からの脱却』日本図書センター。
- 金子元久、2007、「国立大学法人の評価—何が課題か」『IDE 現代の高等教育』490、4-12。
- 金性希・林孝之・齊藤貴浩、2009、「認証評価による大学等の改善効果の創出構造—大学等に対する認証評価の検証アンケート結果の比較分析を中心に」『大学評価・学位研究』9、19-42。
- 木村孟、2011、「大学評価再考」『IDE 現代の高等教育』533、11-15。
- 小玉重夫、2003、『シティズンシップの教育思想』白澤社。
- Kristensen, Bente, 2010, “Has External Quality Assurance Actually Improved Quality in Higher Education Over the Course of 20 Years of the ‘Quality Revolution’?,” *Quality in Higher Education*, 16(2), 153-157.
- 工藤潤、2010、「大学基準協会が実施する新大学評価システム—内部質保証の構築の重要性」『大学評価研究』9、17-27。
- Lewis, Richard, 2005a, 「講演録：ボローニャ宣言—ヨーロッパ高等教育の学位資格と質

- 保証の構造への影響」(吉川裕美子訳)『大学評価・学位研究』3、77-90。
- Lewis, Richard、2005b、「講演録：国際的な教育の質の保証の動向—INQAAHE の活動を中心に」(齊藤貴浩訳)『大学評価・学位研究』3、119-128。
- 前田早苗、2009、「大学の質保証における認証評価が果たすべき役割について」『大学評価研究』8、53-63。
- 宮浦崇・山田勉・鳥居朋子ほか、2011、「大学における内部質保証の実現に向けた取り組み—自己点検・評価活動および教学改善活動の現状と課題」『立命館高等教育研究』11、151-166。
- 村田直樹、2004、「英国高等教育の質保証システムについて—イングランドを中心に」『IDE 現代の高等教育』464、53-61。
- 南島和久、2009、「NPM の展開とその帰結—評価官僚制と統制の多元化」『日本評価研究』9(3)、17-27。
- 南島和久、2010、「NPM をめぐる二つの教義—評価をめぐる『学』と『実務』」山谷清志編『公共部門の評価と管理』晃洋書房。
- 南部広孝、2010、「大学教員の労働と生活」『IDE 現代の高等教育』519、42-46。
- 納谷廣美、2011、「新時代を迎えて—これからの認証評価制度について」『IDE 現代の高等教育』533号、22-27。
- 大場淳、2007、「ボローニャ・プロセスにおける質保証の枠組構築とフランスの対応」『COE 研究シリーズ 28 高等教育の質的保証に関する国際比較研究』広島大学高等教育研究開発センター、45-74。
- 大場淳、2009、「日本における高等教育の市場化」『教育学研究』76(2)、185-196。
- 小笠原道雄編、2003、『教育の哲学』放送大学教育振興会。
- 大南正瑛編、2003、『文献選集 大学評価』エイデル研究所。
- Stensaker, Bjorn and Lee Harvey, 2011, “Accountability: Understanding and Challenges,” Bjorn Stensaker and Lee Harvey, (eds.), *Accountability in Higher Education: Global Perspectives on Trust and Power*, New York: Routledge.
- Schwarz, Stefanie and Don F. Westerheijden, (eds.), 2007, *Accreditation and Evaluation in the European Higher Education Area*, Dordrecht: Springer.
- 館昭、2005、「国際的通用力をもつ大学評価システムの構築—『認証評価』制度の意義と課題」『大学評価・学位研究』3、5-18。
- 館昭、2008、「動き出した認証評価と今後の課題」『IDE 現代の高等教育』476、5-12。
- 田中弥生、2009、「評価可能性のアセスメント (Evaluability Assessment) —大学の自己評価能力向上のために」『大学評価・学位研究』10、27-44。
- 田中弥生、2011、「国立大学法人制度と評価」『シンポジウム—大学における教育研究活動の評価をどう考えるか』報告資料 (2011.12.11、東京理科大学森戸記念館)。
- 遠山敦子、2004、『こう変わる学校 こう変わる大学』講談社。

- 戸澤幾子、2011、「高等教育の評価制度をめぐって—機関別認証評価制度と国立大学法人評価制度を中心に」『レファレンス』61(1)、7-28。
- トロウ・マーチン、1976、『高学歴社会の大学—エリートからマスへ』（天野郁夫・喜多村和之訳）東京大学出版会。
- Trow, Martin, 1996, “Trust, markets and accountability in higher education: A comparative perspective,” *Higher Education Policy*, 9(4): 309-324.
- ユネスコ・APQN、2008、「ユネスコ—APQN ツールキット—国境を越えた教育の質の規制」（齊藤貴浩訳）『大学評価・学位研究』8、67-98。
- Westerheijden, Don F., 2007, “States and Quality of Higher Education,” Don F. Westerheijden, Bjorn Stensaker, Maria Joao Rosa, (eds.), *Quality Assurance in Higher Education: Trends in Regulation, Translation and Transformation*, Dordrecht: Springer-Verlag.
- 横山恵子、2007、「日本型評価国家における私立セクターの特性—機関別認証評価の私立大学への影響に関する研究」『大学論集』38、143-158。

## 第 1 部

# 大学における教育研究活動の評価のシステムに 関する国内外の動向の整理

## 第1章 評価に関する議論の整理と今後の課題

広田 照幸（日本大学）・二宮 祐（一橋大学）

### はじめに——本章のねらいと構成

第1部は、今回の委託調査の中で、「大学における教育研究活動の評価のシステムに関する国内外の動向の整理」を担当した第1グループの研究成果をまとめた論考が並んでいる。第1グループでは、「教育」の質の保証、維持・向上という観点から、大学における教育研究活動の評価のシステムに関して内外の事例を検証しつつ、今後の望ましい評価の在り方について議論するための基礎的な整理を行ってきた。また、また実務的なテーマとして、現在日本学術会議が策定を進めている分野別の教育課程編成上の参照基準が、認証評価の枠組みの中でどのように位置付けられることが望ましいのか等の問題について、ラーニングアウトカムズに着目した内部質保証との関わりも考慮して検討を行った。

本章は、その第1グループの研究を総括的に位置付けるために、以下のような構成となっている。まず前半（第2・第3節）では、大学における教育研究活動の評価のシステムに関する議論のベースにある評価の目的や方法などに関わる考え方を整理する。特定の具体的な制度について丹念に検討するのではなく、さまざまな評価制度の議論の中に含まれる目的や方法などをある程度抽象的なレベルで整理をするのが、そのねらいである。第3節では、誰が何のために評価するのかという問題を扱う。評価の主体や目的が多元的に存在しているから、それらの整理が必要だと思われるからである。第4節では、何をどのように評価するのかという問題を扱う。さまざまな評価手法や技術がもつ可能性と限界とを確認するためである。これらは本報告書の序論でなされている提言をまとめるための基礎的な作業でもあった。

後半（第4節）では、欧米の教育研究評価の現状についてとりあげた研究会での報告の概要と、それらの報告をふまえた研究会の議論とを簡単に紹介する。欧米の評価システムの動向をどうみるのかについて、興味深い論点が登場している中に出てきているので、それを紹介することは意義があるだろうと考えたからである。

### 第1節 誰が何のために評価するのか

#### （1）内部評価と外部評価

本節では、喜多村和之と金子元久の議論を手がかりにして、「誰が何のために評価するのか」という観点から評価に関する議論を整理していきたい。

「誰が」という点でみると、大学の教育研究の評価は、多くの論者が述べているとおり、大きくは2つに分類することができる。内部評価と外部評価である。喜多村和之は、内部



評価を「大学が自らの主導権で自己を評価する行為」、外部評価を「大学の外部者（例：政府）が大学を評価する行為」と説明している（喜多村 1999＝2003, p.228）。

外部評価には、当然のことながら、外部者が大学に対して持つ期待や願望が反映することになる。外部評価には政府ないし非政府機関による評価が考えられるが、「大学評価にとって最も微妙かつ重要な問題は、何といても大学に対する外部評価の要求と大学側の大学自治の主張との間の葛藤であろう」（同論文, p.233）、と喜多村は述べている。

行政権力の下での大学評価は、学問の自由や大学自治の侵害の見地から望ましくないばかりでなく、基本的には教育・研究という、自発的で効率性という尺度になじまない機能の評価には、適切でない場合が少なくない。とりわけ政府が大学評価に介入してくる場合には、ときの政治的、財政的状況の影響を受けがちであるのみならず、画一的な評価尺度を硬直的に大学に適用される恐れ少なしとしない（同論文, pp.233-234）

ここで喜多村は、大学における教育・研究が自発的なものであり、かつ効率性という尺度になじまないものとしてとらえることで、外部評価の要求と大学側の大学自治の主張との間の緊張関係と、外部評価のもつ危うさを指摘している<sup>1)</sup>。「大学自治」を古びた時代遅れの理念とみなす向きもあるが、「評価の氾濫」が懸念される状況の中では、この喜多村の視点は重要である。大学が持続的に知の生産や革新に貢献し、またその教育によって生まれる人材が不透明な現代社会に対して長期的に貢献的な役割を担っていくためには、大学は一定程度、社会の即時的なニーズから距離をとる必要がある。「大学自治」は中世大学の遺物などではなく、大学が短期的・即時的な社会のニーズと一線を画すために必要な制度・慣行であると考えられることができるだろう。

大学教育の「質」や研究の「水準」を簡単な尺度で測ることが難しいのは、一つには、その判別には一定レベルの学問的な専門性や大学教育についての理解が必要だからであり、もう一つには、教育が本源的に持つ不確定性（広田 2009）や、個別の研究成果が知的な発展に対して持っている意義の説明しがたさに由来しているだろう。そうであるがゆえに、外部から持ち込まれる評価の項目や尺度は、しばしば、大学の教育研究の実情とはずれたものが押しつけられるような印象を、大学内部の者に与えることになる。

この点は、ヨーロッパでも同様である。高等教育における「質」をめぐる議論の流行と経営思想との関わりを考察したステンセイカー（Stensaker 2007）は、大学外部の者が望むほど改革が進まないのは、大学が単に怠惰だからなのではない、と論じている。技術的に有効でないもの、役に立たない改革が採用されるために、それに抵抗する面がある一方で、自分たちのあり方を見直す契機になる場合もある、と。評価のあり方は、この議論と重なっている。技術的に有効でない評価や、教育研究の改善・向上に役に立たない評価は、いくら外部の者にわかりやすいものであったとしても、現場の教員にとって受け入れられるようなものではない。

ヨーロッパの質保証の仕組みの展開を批判的に振り返るニュートンは、1990年代初めの外部から大学に対して押しつけられた改革が、結果的にモラル・ハザードを生み出した、と論じている。質保証を売り込むさまざまなアクターが、各々の「質」の定義を売り込み、彼らはそれが現実に質の向上につながるのか、学生の経験が良くなるのかに言及しなかった（Newton 2010, pp.51-52）。それゆえ、当の大学の教員たちは、「質」の強調をまったく異なる意味のものとして経験することになった。それは、実質的な大学教育の改善向上とは無縁な、重荷（burden）、官僚制（bureaucracy）、ゲームプレイング（game playing）として経験されている、というのである（Newton 2002, Newton 2010）。こうした状況は、多かれ少なかれ、近年の日本の状況でもあるといえる。

もちろん、ステンセイカーがいうように、外部からの評価の項目や尺度が自分たちの自明視してきた慣行や実践を見直す契機になる場合もあるだろうから、外部評価がいけないというわけではなく、それは両義的なものである。少なくとも、「大学内部の者、大学関係者は、既得権や現状維持を求めている信用ができない。外部者こそ、本当に大学が必要なものが何かを知っている」というような一方的な見方ではなく、もっとニュートラルに、喜多村が述べるような外部評価の要求と大学側の大学自治の主張との間の緊張関係としてとらえる必要があるだろう。

## （2）改善・向上のための評価と判別のための評価

「何のための評価か」という点で大きく分けると、改善・向上のための評価と判別のための評価という区分が可能である。教育評価論の中では、教育過程へと情報をフィードバックして教育指導の改善を目的とした「形成的評価」と、完結した教育過程の成果を外部に可視化することを目的とした「総括的評価」という区分で語られることが多い。両者は、情報の性質のみならず、情報の作られ方や扱われ方がまったく異なっている。

機関や政策の評価においても、改善・向上のための評価と判別のための評価は、まったく異なる性格や機能を持っている。評価の志向性という点から大学評価には2つの理念があると述べる金子は、その二つを次のように論じている。

ひとつは、微視的に組織的にも、よりよい高等教育の機能を発揮するために、様々な手段を統制（コントロール）するための情報を得る手段として評価が位置づけられることである。これを「統制」の原理とっておこう。これに対して、いま一つは「自主性」の原理ともいうべきものである。これは評価の対象自体が、評価の基準に強く関わることに関係している。これら二つの原理を如何に調和させるかが、大学評価の具体的な形態の基本的な課題であるといってもよからう（金子 1991=2003, p.127）

評価の対象者自身が関わる内在的な基準から、自らの活動の改善・向上（育成）をめざして行う評価と、外部者（他者）が外在的な基準で高等教育の質を判別する

ために行う評価という2つである（図1）。

|    |     |        |
|----|-----|--------|
|    | 自主性 | 統制     |
| 主体 | 自己  | —— 他者  |
| 目的 | 育成  | —— 判別  |
| 基準 | 内在的 | —— 外在的 |

出典：金子（1993=2003, p.127）

図1 評価の志向性という観点からの概念化

金子の整理はとてもすっきりしているが、やや単純すぎる。というのも、「主体」「目的」「基準」がこの組合せ以外にも考えられるからである。すなわち、内在的な基準は、前述したように、外部者（他者）にはなかなか設定がしにくいいため、「主体」と「基準」とはセットとして考えられるが、育成—判別という「目的」は、「主体」「基準」とは別の組合せが考えられるのではないだろうか。たとえば、「自己—判別—内在的」という組合せの例を考えてみるができる。大学執行部が雇用している教員を査定（＝判別）したり、教員同士が不出来な同僚を排斥（＝判別）したりするために評価を行うといったケース（大学教員とすると憂鬱になるが）がそれにあたる。あるいは対照的に、「他者—育成—外在的」という組合せを考えることもできる。それは、大学の外部の評価機関などが、それぞれの大学の教育研究の改善・向上（育成）を目的として外から持ち込んだ尺度や手法で評価を行う、といったケースなどである。

そこで、ここでは、「目的」の部分に注目し、改善・向上（金子のいう「育成」）のための評価と判別のための評価という2種類の評価について、もう少し考えてみよう。

改善・向上のための評価と判別のための評価という2種類の評価は、ともに現状の問題点を明らかにし、改革努力を刺激するという点ではポジティブな効果をもつけれども、同時に、まったく異なる種類の問題を伴ってもある。

### （3）改善・向上のための評価が抱える難点

教育指導における形成的評価と、機関や政策を対象にした改善・向上のための評価とは、前述したとおりよく似ている側面がある。しかしながら、決定的に違うのは、前者は指導—被指導という教育的な関係（教師—生徒関係）があらかじめ制度化されており、それを前提とした評価であるのに対して、機関や政策を対象にした改善・向上のための評価については、その関係の構築自体が論争的な課題になるということである。換言すれば、このタイプの評価の制度化は新たな指導—被指導関係を作り出し、既存の社会諸関係を変容させるということである。

ここから改善・向上のための評価が抱える難点の多くが出てくる。

第一に、評価の制度化は強い統制機能をもっているということである。たとえば、政府

や第三者機関が大学に対して評価を行い、改善・向上のための勧告や指示を出す制度があったらに作られたとしたら、それは、政府や中間団体による大学の質的な統制という関係ができたことを意味する。同様に、大学の管理者が学部や学科に対して、さらには、学部や学科の評価セクションが個々の学科や教員に対して評価をおこなう仕組みが制度化されることは、大学組織内の統制—被統制関係に新しいものがつけ加わったことを意味している。それゆえ、自治や自律性を伝統的に重んじてきた大学という組織では、それゆえ、改善・向上を名目としたものであっても評価に対する抵抗感を強くもつのは当然である。「評価を通した統制」を警戒するからである。

第二に、教育研究についての考え方のちがいがコンフリクトを生む。「何が改善・向上なのか」について、ある特定の考え方が強調される時、それとは異なる考え方をする者は、改革に向けた議論を拒否したり、抵抗したりする。教育条件の整備や研究資金の増額のようなインフラ的な部分を除くと、大学の教育研究のあるべき姿をどう考えるかについて、誰もが合意するものは存在しない。そのような大学観・大学教育観などのちがいによって、「何が改善・向上なのか」についての考え方も一致しないのが普通である。また、分野や専門によるちがい、カリキュラム上で担当する科目群の違いなどによっても、望ましい方向が異なっていることもある。誰かの目からみて「改善」と思われるものが、別の者には「改悪」として映る。

第三に、改善・向上のための評価が外形的で形式的な尺度を含む場合には、その尺度から何かがこぼれているという感じを、個々の部局や個々の教員が抱くことも起きる。「何が改善・向上なのか」を外部の者が定義し、ごくかぎられた特定の尺度を使って部局間や教員間、授業間の比較を行うとすると、それは当事者の不満やみかけだけの取り繕いを生んだりしてしまう。

このように考えると、大学外部の者が望むほど改革が進まないのは、大学が単に怠惰だからなのではないという、ステンセイカー（Stensaker 2007）の議論をもう一步踏み込んで考えることができる。大学という組織のもつ特徴——組織内の多元性や大学観・大学教育観の多様性、個々の部局・教員の自律性など——を無視して、ある方向の「改善・向上」を外から押しつけることの困難さや問題なのである。

もちろん、評価を通した改善・向上がどういう方向になるかはわからないが、それぞれの大学、それぞれの部局は、何らかの改善や向上が可能なはずである。もしもそれを大学自らが自主的に追求していくのであれば、それは、上で述べた、大学という組織に固有の特徴である多元性・多様性・自律性のある程度まで尊重して、その強みを生かしたものでなければならないだろう。

その際の一つの方法として、共通の目的・目標を共有するための議論の過程自体を重視する、というやり方がある。考え方が対立する複数の部局の間や、教員間での大学観・大学教育観が異なる教員間で、評価を介したコミュニケーションがなされること自体の意義に注目をあてるということである。教育研究の現状や教育研究が目指すべきものについて

組織の成員が率直に話し合い、相互の共通点と相違点を確認することが、自発的な改善・向上の努力や目標を引き出すことになる。そういう取り組みがなされる場合には、改善・向上のための評価は、議論のための契機や素材として重要な役割を果たすことになるだろう。

#### **（４）判別のための評価が抱える難点**

外在的な基準で、判別を目的としてなされる評価は、大学の教育研究を対象として、何をどこまでやれるのだろうか。理論上・形式上は多様なやり方が考えられるし、制度化されても来ているが、ここにもいくつかの重要な難点が存在する。

第一に、多くの指摘がくり返されてきたように、評価において外在的な基準として使用される指標・尺度の恣意性の問題である。

たとえば、複数の尺度を合成して総括的な評価を数値で示すような評価の場合、個々の尺度の妥当性や信頼性の問題だけでなく、何に関して尺度を選定するのかという点やそれぞれの項目にどういうウェイトをかけて総括評価を行うのかといった点もまた、恣意的なものであることを免れない。大学のランキングは典型的な例である。金子（2007a）は、研究能力の指標の乱暴さや、「評判」のウェイトの大きさが、既存の暗黙の秩序を強化する機能を持つなど、ランキングとしての指標としての問題点を列挙している。

また、国立大学法人評価における目標設定のように、それぞれが設定した独自の目標の達成度を測るような評価が行われる場合には、評価結果自体が相対的なものなので、相互の比較や序列付けはできない。それゆえ、それぞれが設定した目標の達成へのインセンティブを作り出す装置としての評価としては機能しうるけれども、資源配分などのための情報として利用するのは適切ではない。高い目標を掲げて十分達成できなかったA大学と低い目標を掲げて達成したB大学とを相互に比較することは、原理的にはできない。たとえば、国立大学法人の中間評価の仕組みを考察した金子（2007b）は、「教育研究の評価に関して、『関係者の期待』が基準とされ、しかもそれは大学の特性や条件によって異なることになっているから、大学の側が想定する水準と機構の側のそれとが食い違うことは十分にありえる」と指摘している。独自の目標を掲げるかぎり評価の基準は相対的なものでしかありえないから、誰もが合意できるような達成度の評価は原理的にできないと考えるべきであろう。

つまり「判別」はあいまいさや恣意性を免れないのである。もちろん、指標や尺度の恣意性を理解した上で、ランキングを詳細に分析検討して大学や部局がそれに向けた戦略を立てるということはある（Taylor and Braddock 2007, 坂本 2007）。ある大学や部局があるユニークな目標を戦略的に設定して、数年間の間の組織成員の注意や努力をそこに傾注させていく手段として活用する、ということもできるだろう。だがそれはあくまでも、個々の大学や部局が周到な戦略のもとで評価を活用するという意味での効用であって、大学ランキングやそれぞれが設定した独自の目標の達成度を測るような評価では、外部に発

信されるわかりやすい数値での評価結果は、指標・尺度の恣意性があるため、情報としての利用法は謙抑的なものでなければならない。朝日新聞社で『大学ランキング』の編集に関わる小林哲夫（2007）は、あえて総合ランキングを作っていないという。それは、ここで述べたような問題点をよく理解しているからであろう。

第二に、外在的な基準で判別を目的としてなされる評価は、にもかかわらず、情報の非対称性を克服できない。つまり、外部者に評価結果を示すことが、大学内の成員と同様にその大学を知る手立てになるわけではない、ということである。

判別のためになされる評価は、通常、二種類のタイプの情報として外部に発信される。一つは、項目別であれ、総括的なものであれ、序列や点数評価など、結果を縮約した情報にまとめたものである。もう一つは、評価に関連した大学内の情報をそのまま、できるだけ多く外部に公開させるというものである。しかし、前者は、縮約された情報であるだけに、大学の教育研究の内実に関する具体的な情報を含まない。それゆえ、ランキングや序列付けなどには使うことができるが、その大学の教育研究の実際を知る手がかりにはならないのである（A大学が THES の国際ランキングで何番目かということは、ある分野で学ぶためにその大学に入学すべきかどうかを決める情報にはならない）。後者は、逆に、詳細な情報の公開を求めれば求めるほど、学部者には何が有益な情報なのか取捨選択できなくなる。分厚すぎて誰にも読まれない評価報告書だけが積み上がっていくことになる。

この問題は、判別のための評価を誰がどういう情報によって行うかという点で2つに分けることができる。一つは、認証評価機関やランキング作成機関など、外部の何らかの機関が判別するようなやり方である。もう一つは、さまざまな学内情報をそのまま公開させて、直接外部の消費者（受験生や企業など）からアクセス可能にすることで、彼らに判別（評価）を委ねてしまうやり方である。

前者の場合、大学から出てくる情報が少なければ判別（評価のとりまとめ）が不可能だが、逆に、膨大な情報が出てきたとしても判別（評価のとりまとめ）が困難になる。かぎられた人員と時間とで行う評価には、途方もない量の情報を与えられたとしても、それらを適切に生かした評価をおこなえるわけではないのである。また、外部の一般の消費者にとっては、ランキングや序列はわかりやすく好まれるが、これは前に述べたとおりあいまいさや恣意性を帯びたものにすぎない。だからといって、膨大な学内情報が公開されていたとしても、それを丁寧に閲覧し、分析的に考察するスキルが欠けている（熱意や時間もかけているのが一般的である）。

そもそも、大学の側の問題を解決しないといけないということ以上に、現代の大学に対する世間の無理解や過剰な不信感を改善していく手立てが必要なのではないだろうか。この点については、ここでは書ききれないが、たとえば、企業と大学の活動の性質の違いが十分理解されていない。また、大学教員自身が学生の現状を過度にネガティブに語りがちだし、大学の外の人たちは、30 年前、40 年前の自分の体験に依拠して「今」の大学教育の問題を論じたり、問題を抱えた逸話的な事例を過度に一般化して論じてしまう傾向もあ

るだろう。

葛城浩一（2011）によれば、ボーダーフリー大学の若手教員は、研究だけでなく教育にも強い熱意を持ち、また所属する組織に対するコミットメントも高い。そんな教員たちの実情が世間に知られていない。そして、改革の努力の跡や一生懸命やった挑戦的な試みの資料には目も向けず、ランキングや序列のような簡単で相対的な位置を示す数値のみをほしがっている。そうした、世間の人たちが抱く大学の実情に対する根強い思い込みや、発信される情報に対するリテラシーの欠如が、判別のためになされる評価の拡大を支えているとするならば、いくら評価結果の公表や大学内部の情報の公開を進めていったとしても、事態は改善され得ない。「評価疲れ」「評価の水ぶくれ」が問題になってきていることを考えると、世間の人たちの偏見を正し、大学の教育研究の現状に対して正しい認識を持ってもらうことこそが、何よりも喫緊に取り組むべきことではないだろうか。

### 第3節 何をどのように評価するのか

#### （1）評価基準の種類——3つの基準、2つのアウトカム

大学教育の質をどう評価についても複数の考え方がある。ここではまず、金子元久（2011）の整理をみておこう。

金子は、大学教育の質を、どのような基準で、またどのような形で「保証」するのかについて3つのタイプを挙げている。①絶対基準、②インプット基準、③アウトカム基準、の3つである。評価において使用されるのは、この3種類の基準のどれかである。

第一の「絶対基準」は、「大卒者の知識・能力が、明確に定義された一定の水準を満たしていること」である。しかし、金子は、こうした考え方による質保証は現代では困難になっている、という。というのも、一つには、一部の専門職分野を例外として「卒業生に要求される知識技能が、必ずしも厳格に定義しえない」からである。もう一つには、「知識・技能には多様な側面があり、それをかぎられた試験のみで判定することは難しい」（金子2011, p.6）からである。

第二の「インプット基準」は、「大学における教育内容について、明確に基準を定め、それを満たしていることを基準として、質の保証とみる考え方」である。中等教育の延長上に発展したアメリカの大学は、こうした考え方をもとに適格認定制度を作り上げた。「授業科目の構成、個々の授業（単位）の修得に必要とされる授業時間数と学習時間数、教員の数と構成、諸施設、さらにその運営、といった要素が完備される大学が学位を授与することができる大学として認定される」。それに対して、日本では戦後導入された適格認定制度が質保証の役割を果たすことができなかったため、「政府が設置の際に一定の基準が備えられていることを基準として認可をおこなう設置認可制度が実質的に質保証の役割を果たしてきたことは周知の通りである」（同論文, p.7）<sup>2)</sup>。

金子は、そこには問題があった、という。教育の中身に踏み込まないで「教員数、設備

など外形的なインプットのみを基準として質保証が行われるという傾向がもたらされた」  
うえ、社会的な要求に対応していない、という点である。「最も基本的な問題は、インプット基準が、学生が何を学習するかという視点を十分に踏まえたものではない、という点である」(同論文, p.7)。

第三の「アウトカム基準」は、「大学教育が、学生にどのような知識・能力を形成したかを基準とする質保証である」。ここでいうアウトカム基準の質保証は、第1の絶対基準のそれとは異なるものである。「絶対基準の質保証では、卒業者の知識・能力そのものについて一定の最低基準を設定することが想定されていた。これに対してアウトカム基準が着目するのは、大学教育が、学生が一定の知識・能力を形成するうえで、どのような効果があるのか、という点にある」(同論文, p.8)。

近年の大学評価の流れは、欧米でも日本でも「インプットからアウトカム重視へ」というものである。しかしながら、「アウトカム重視」という議論には、2つの種類の評価のあり方が存在している。一つは、「アウトカムの測定をもとにした評価」である。もう一つは、「アウトカムを明確にした教育改善 (の評価)」である。この2つをきちんと区別しないと、議論が混乱することになる。両者はともに欧米で議論され、実践されているが、前者はまだ部分的・試行的なものにとどまっている。現実的な形で制度全体をカバーするような議論や実践は、後者のものであるように思われる。以下では、この2つについて順番に考えていくことにしたい。

## **(2) アウトカムの測定**

まずは「アウトカムの測定をもとにした評価」である。たとえば川嶋 (2009) のように、個々の学生の学習の成果をテストや試験によって直接測定し、それを教育の成果として評価に用いるものである。議論や実践を見ていくと、テストや試験の利用のされ方の違いによって、それらはさらに2種類の評価方法を区別することができる。

### **1) 個別化した利用法**

一つは、学生個々の判定を含んだ個別性にもとづくツールである。つまり、測られた当人の評価、あるいは、当人が所属する集団の評価にテストや試験の結果が用いられるという使い方である。もっと良い用語があるのかもしれないが、ここでは「個別化した利用法」と読んでおく。この利用法は、個人または集団を判別することを目的としたものである。

たとえば、特定分野の知識やスキルを測って選抜や資格付与に使う、というアウトカム測定がある。医師や法曹などの国家試験、特定の分野や特定の資格に関して行われている各種の検定試験や資格試験などがそうした例である。当の学生の選抜や資格付与だけでなく、場合によっては、大学別に合格者数や合格者率などを算出して、大学間で成果を比較する、といった使われ方もする。ただし、分野によっては不当にカリキュラムがしばられてしまい、教育内容の画一化や硬直化を生んでしまいかねない。また、「試験のために学習



をする」という行為を呼び込んでしまいかねず、大学における知の創造性や柔軟性に悪影響を与えることもありうる。

また、近年急速に広がっているのは、ジェネリックな知識やスキルのテストである。羽田貴史（2012）が指摘しているように、米国ではかなりの大学が認証評価の際のデータとして採用するようになっているし、日本では企業が採用希望の応募学生を一時選抜するための道具として流行している（SPI など）。しかしながら、この種のテストが本当に教育のアウトカムを測っているのかという点で疑問がある。大学入学以前の学生の「素質の差」によるところが大きいのではないかという指摘もある。反復して測定しないかぎり「伸び」を測ることはできないし、逆に（SPI がそうだが）集中的な練習でスコアを変動させることもできる。さらに、大学のほとんどの科目は高度な教養的知識や特定の専門的知識を伝達することを主要な役割としており、ジェネリックなスキルの涵養を直接的な目的とはしていないので、教育の成果としてはやや筋違いものを測っているともいえる。もしも高度な教養的知識や深い専門的知識をまったく教育しないで、ひたすら体験学習などでジェネリックなスキルを形成する教育機関があったとしたら、それは「大学教育」の名に値するかどうか、疑念が表明されるはずである。

金子はアウトカム評価の技術的な問題点として、適切性・妥当性への疑問、実施可能性・信頼性への疑問を提示している（金子 2009a）。その通りである。また、学習成果の数値自体からは、その成果をもたらした要因を特定することができない（羽田 2004）。学生間、機関間、プログラム間の比較はできるとしても、だからといって具体的な改善のための情報としてはあまり役立たないのが現実なのである。

ジェネリックな知識やスキルのテスト自体が無意味なわけではないが、少なくとも、個々の科目や各回の授業のレベルでジェネリックな知識やスキルを用いた効果を検証しようとしたりするのは、あまり意味がないことであろう。

## 2）一般的知見を引き出す情報としての利用法

もう一つの「アウトカムの測定をもとにした評価」の方法は、さまざまな特性や経歴を持った個人の間でのアウトカムのちがいを把握することで、環境や教育の諸条件がアウトカムにどう影響を与えているのかを一般的知見として得ようとする利用法である。ここでは、「一般的知見を引き出す情報としての利用法」と呼んでおこう。この利用法は、改善・向上を目的としたものである。

金子（2011）は、アウトカム測定を三種類に分けている。①専門領域別の学習達成度テスト（College Basic Academic Subjects Examination (College BASE)、Major Field Test (MFTs) など）、②一般的な能力（Generic Skill）を測定するテスト（Collegiate Learning Assessment (CLA) など）、及び、③学生の意識や行動についての質問紙調査（全米学生学習行動調査：National Survey on Student Engagement (NSSE) など）である。この分類に依拠していえば、①と②は主として前項の個別化した利用法に用いられているが、

③は主として一般的知見を引き出す情報として利用されている。

この利用法は、社会学や心理学の統計的な手法を用いてサンプル内でのグループやカテゴリー間を比較する形で、アウトカムの差異を説明しようとする。たとえば、特定の属性や経歴（例：高校時代に勉強時間が長かったグループとそうでないグループ）を独立変数とし、アウトカム（例：大学での学習時間）を従属変数とし、相互の関連がどの程度か、また、カリキュラムのあり方や生活スタイルのちがいによってどの程度その差が影響を受けるのか、あるいは、大学4年間の間にそれがどのように変容したのか、といったことを考察する。

こうした利用の仕方から得られるのは、アウトカムの高低を左右する属性要因や環境要因についての知見である。単純な言い方をすれば、どういう学生が、どういう環境の下で、どういう教育を受けていれば、高いアウトカムを示すのか、ということが明らかにされる（山田 2009, 山田 2011, 金子 2009b）。研究としてなされるこうした考察は、サンプルとなった大学や学生の教育のあり方を改善する手がかりを与えるだけでなく、似たような条件や状況の下にある大学や学生の教育の改善のための知見となることができる。その意味で、この利用法は、ごく少数のサンプル調査によって広範な適用性を持った知見を引き出すことができるメリットがある。

とはいえ、このやり方にはいくつかの限界がある。

第一に、誰もがやれるわけではない。一つには、調査票の設計や変数の同定、データの処理や解釈などに専門性が必要だからである。質問紙調査法のスキルに熟達しない者が見よう見まねでやっても、決して信頼できるような知見は得られない（出来の悪い実態調査にとどまる）。また、調査票の設計から分析までかなりの手間と費用がかかる。しばしばすでに実感としてわかっていることが数字で裏づけられるぐらいの結果になるから、徒労感も生まれてしまうだろう。だから、一部で主張されているような、すべての大学が自前でそのような調査分析やることを求めるのは困難だし、ムダである。限定されたサンプルデータを用い、きちんとした専門家によって入念に設計されたデザインで行われる調査から導き出された一般的知見を積み上げていき、それを各大学で参照しつつ有効活用するほうが、明らかに効率的である。

第二に、回答者のモチベーションや戦略的対応の問題がつきまとっている。米国の大学で認証評価のためのデータとするために、学生に対する調査がしきりに行われているが、困難の一つは、回答者に熱心に取り組ませることの難しさである（星・鈴木 2009, 金子 2009a）。当人の評価に関わるアウトカム測定と異なり、匿名性を保持したままの調査の場合には、モチベーションの調達が大きな課題となる。米国の CLA（College Learning Assessment）実施校のヒアリング調査を行った星千枝らによれば、「受検料を大学が負担する上、受検謝礼を学生に支払って CLA を実施していた。このような実施方法では、受検者が確保できたとしても、学生が真面目に取り組むかどうか疑問があ」った（星・鈴木 2009, p.55）。だからといって、記名などで個人を特定する形でデータを収集するならば、

回答者は戦略的に回答してしまう（普段本をほとんど読まない者が「よく読む」に答える、など）。大学名を明らかにしてアウトカムを比較する調査をやったとしたら、どの大学も調査の実施過程でさまざまな工夫（望ましい回答のほめかし、など）をしてしまうかもしれない。つまり、あくまでも匿名性が保持されないとデータが歪んでしまうのだが、だからといってその形式では、回答者のモチベーションの面で問題が残る、ということである<sup>3)</sup>。

一般的知見を引き出す情報としての利用法については、もう一つ大きな問題がある。個人を匿名化して集団レベルでアウトカムがどれぐらいかを測って比較してみたとしても、その結果は個々の学生の選抜や指導の情報としては価値が低い、ということである。東京大学の卒業生が全体として高いアウトカムを示したとしても、それは集合的な情報にすぎない。その大学のある卒業生の採用を考慮する企業は、結局、別のツールであらためて個人の能力・適性やスキルを測ることになる。また、演習への出席率が低い学生がジェネリック・スキルの伸長の程度が低いという一般的な傾向が得られたとしても、ある学生が演習への出席率が低いからといって、ジェネリック・スキルの低さを前提に指導するとすると、まちがってしまうケースが続出するだろう。アウトカム指標を取り入れたとしても、それによって大学は個々の学生・卒業生の「質」を保証できるわけではないのである。妙な期待を抱いて、アウトカムを測る努力を大学が積み重ねたとしても、結局は個々の卒業生の採用時の選抜はまた別の基準や手段でなされてしまうことになるだろう。

### （3）アウトカムの明確化による教育改善とその評価

アウトカムを直接測定するというのではないもう一つの「アウトカム重視」のやり方は、各教育機関や部局が「どういうアウトカムをめざして教育していくのか」を明確にして、それによって、カリキュラムや授業の改善工夫を行い、それが適切になされているかどうかを外部者が評価する、というふうなやり方である。

近年の評価のあり方をめぐる大きな流れは、どちらかというところの方向が盛んである。1990年代に詳細な分野別評価を行っていた英国は、評価に関わる負担や効率を顧慮して、2000年代には分野別参照基準（Subject Benchmark Statement）を作成して、各大学がそれを活用してカリキュラムや授業の編成原理を説明することを課す、というやり方に切り替えた。こうした英国のQAAの仕組みを参考にしながら日本学術会議が提案しているスキームもこれである（日本学術会議 2010）。大陸ヨーロッパ諸国でも、ENQAが2005年にまとめたガイドラインは、同様に、各国の高等教育機関が自らの教育の質を説明できることを重視するスキームを採用している（Clark 2007）。福留東土（2005）によれば、米国のア krediteーションにおけるアウトカム評価もまた、学生のパフォーマンスの測定ではなく、各大学がアウトカム像を明確化し、その像の実現に向けた教育をいかに改善しているのかを外に向かって説明するという点が重視されている。

もちろん、こうしたやり方は、必ずしもアウトカムの測定自体を不必要にするものでは

ない。教育の現状にどういう問題点があるのか、あるいは、教育の改善工夫がどのような結果を生んでいるのかをモニターする手だてとして、アウトカムの測定結果が用いられることもある。しかしながら、このやり方で最も重要なのは、アウトカムの測定結果それ自体ではなく、それに直接・間接に影響を与えているであろう教育課程や教育方法の編制のされ方なのである。江原武一（1994=2003, pp.72-76）がいうように、評価に用いられる指標には、(1)主観評価、(2)量的指標、(3)評価尺度、(4)記述的個別分析の4つがありうる。どの指標もそれぞれ長所・短所があるけれども、少なくとも、一般に「アウトカムの測定」でイメージされるような評価尺度による指標は、仮に利用されとしても、ここでは評価のための多様な参考情報の一つという位置付けにすぎない。

ここまではかなり一般的なレベルで評価の目的や方法について考察してきた。次節では、欧州と米国の教育研究の評価の動向について、研究会でなされた報告と議論を整理しておく。

## 第4節 欧米における教育研究の評価の動向

### (1) 欧州各国における研究評価

この項では、2011年7月2日の研究会において、大学評価・学位授与機構の林隆之氏から報告していただいた「大学の研究評価制度の現状と課題」の概要を紹介し、それをふまえた議論で出た論点を整理する。

林氏の報告の要点は次の通りである。まず言えるのは、欧州各国では研究評価をめぐる動きは流動的な状況にあるということである。単線的に評価システムが改善しているというわけでは決してなく、さまざまな批判に対してアドホックに対応している。現在では、①「評価の弊害、コストが顕在化」が問題視されていて、たとえば、評価の観点に合わせるために研究の質を犠牲にして量を追及すること、評価に関する作業負担の重さが指摘されている。また、欧州の特異な状況として、②政府と大学との関係が変容しつつある。関係者が多様になるにつれて、評価方法と評価結果の情報についての意見対立が生じるようになってきているのである。さらに、③研究活動の多様化に即した評価への転換が主張されている。たとえば、研究者の自主的な活動が自然に社会経済的利益をもたらすという観点ではなく、しっかりと社会的便益まで射程に入れる必要性が主張されている。

そもそも研究評価の困難は、評価にあたって複数の目的がトレードオフの関係にあることである。①研究活動へのブロックグラントの傾斜配分にあるのか、②研究活動の改善促進にあるのか、③アカウンタビリティにあるのか、——それが明確でないと流動的な状況は収まらない。このトレードオフに対して、評価のあり方をどう考えるかが問われる。「評価コストの抑制のために評価を特定の目的に限定する」のか、あるいは、「評価に複数の目的の実現を期待する」のかによって、目指すべき評価のイメージは変わってくるからである。また、評価の基準についても複数のものがある。①大学を比較しての卓越性に重きを

置くか、または、②それぞれの大学の目標に沿ったかたちでの多様性に重きを置くかによって、目指される評価は異なる。ただし、研究成果については前者の「卓越性」を重視する流れがあるようである。

イギリスの例を紹介する。サッチャー政権の頃から評価結果に基づいてブロックグラントを傾斜配分してきたが、2001年のRAEの後、見直しが行われた。そこで指摘されたのは、評点がインフレしていくために資金配分のメカニズムとしてうまく働かないこと、各大学がこのいわば「ゲーム」に慣れてしまうことであった。たとえば、評価に対応するための教員の引き抜きというケースもあったようである。また、評価のコストパフォーマンスが必ずしも良くないこと、機関の枠を超えた共同研究や学際的な研究に対して公平性を保つことができているかどうか分からないこと、そうした問題が主張された。そこで、2014年に実施が予定されている評価では、さまざまな紆余曲折の末に、成果の質、インパクト、環境を評価の指標とすることになった。成果の質は、ビブリオメトリクス指標をパネルが参考にする「インフォームド・ピアレビュー」として、それまで提起されていた、あくまでも指標にくくる観点は後退した。また、インパクトは、経済、社会、政策、文化、生活の質への影響を含めたきわめて広い定義になっていて、大学によって提出されたケーススタディ等を評価することになっている。

以上の林氏の報告をもとにして、第1グループの参加者のディスカッションでは次のような論点が出された。

第一に、日本における高等教育行政の枠内で、研究評価を資金配分に結び付ける必要性の有無や程度に関する問題である。研究評価について科学技術行政・学術行政と高等教育行政とをどのように切り分けるべきかについて、まだ十分な整理が行われているとはいいがたい状況にあります。また、ブロックグラントを競争的に配分する必要性の有無についても議論が必要である。現状の運営費交付金は研究費と教育費に分かれていない。それに対して英国では、研究費と教育費を分けたうえで、研究費を競争的に配るかどうかという議論がなされている。日本ではさらに、法人評価の結果と運営費交付金の関係もわかりにくい状況にある。

第二の論点は、インフォームド・ピアレビューに関するものである。イギリスにおいては、ピアレビューの方法論的妥当性への信頼の揺らぎから、インフォームド・ピアレビューのあり方が模索されてきた。一方、日本においては、そもそもピアレビューの参考となるデータが少ないという状況下で、まずはデータの公開や標準化の必要性が提起されてきた。この違いをどう考えるか。また、個別教員の評価の総和を大学の評価としてしまってもよいのか、目先の研究ばかり優先される、定量的評価ばかりが過大視されてしまう危うさはないのか、といった点も、今後の日本の研究評価のあり方として議論されねばならない重要なポイントである。

## （２）欧州各国における教育評価

次に、2011 年 8 月 6 日の大場淳氏の報告「欧州各国における教育評価」の要点と、それをふまえたディスカッションで出た論点をまとめる。大場氏の報告の要点は次の通りである。

近年、欧州各国の質保証の重点は外部質保証から内部質保証へと移行している。内部質保証は学内の仕組みであって、機関やプログラムに関する質の監視・向上を目的としている。教育課程について内部質保証の事例を挙げると、内容・教授法・達成すべき学習成果の定期的評価、教職員や学生の討論などのような非公式な形による恒常的評価、教育課程設計の効果性、総合性などに関する手続きの定期的評価といったものがある。

質保証の手法をどうするかという問題については、適格認定を意味するアクレディテーション（*accreditation*）が主流ではあるものの、機関を拘束してしまうアクレディテーションに対する大学側の拒否感は依然として強い。そして、まだごく一部の国ではあるが、機関の自律性を前提とした手順審査であるオーディット（*audit*）へと移行しているという状況がある。ただ、もちろん、アクレディテーションの概念は幅広いので、国によってはオーディットに近いものも含まれていることもある。

事例としては、欧州大学協会とフィンランドを紹介する。欧州大学協会の提起する枠組みは、機関評価オーディットによるもので、機関の戦略的経営と内部質保証文化の継続的発達を促すことを目的としている。自己評価に基づいて訪問調査を行って、最後に勧告を出している。これまでに 45 カ国の約 260 機関が受審している。

フィンランドもまたオーディットを実施している。そこでは、機関の自律性が重視されていて、機関みずからが設定した質に関する目標を明確にすること、教育の質の維持・向上のための手法、手順について評価すること、質保証システムが意図されたとおりの機能していて効果的な改善手段となり得ているかを評価をすることといった、ことがそこには含まれている。

こうした事例が出てくる契機の一つとして、欧州において政府と大学との関係がいかにあるべきかという点に関する手探りの動向が背景にあった。伝統的な理解としては、大学は公共財であり、事前規制による政府統制が当然であった。ところが、1980 年代以降、新公共経営論、ニュー・パブリック・マネジメント論に基づいてその枠組みが改革の対象になった。規制緩和が進められ、大学の自律性の拡大が図られたのである。同時に、顧客重視、アカウンタビリティといったことが求められた。一方、独立した評価機関などによる事後的な評価制度の枠組みが整えられ、市場化が図られつつも大学への直接的・間接的な政府統制もまた並存する、という状況になっていった。

また、1990 年代以降の欧州の特徴としては、ステークホルダーが増加したことを指摘できる。国家の上に立つことになる超国家的組織も、あるいは、また国家の下に立つ地方も分権化によってステークホルダーとして力を持つようになり、さらには、「顧客」の一つである学生も当然のことながらステークホルダーとして重視されるようになってきたのであ

る。

以上の大場氏の報告をもとにして、第1グループの参加者のディスカッションでは次のような論点が出された。

まず第一に、欧州の政策背景をおさえておく必要があるということである。前提として、学士、修士、博士という共通の学位課程をつくろうとするボローニャ・プロセスがあつて、そこでは、国を超えて相互の認証を保証するための質保証、および、対外的な適格性の証明としての質保証が必要になってきたということがある。また、ア krediteーションとオーディットの関係も国によって異なっていて、学士や修士を付与するためのプログラムを対象にしたア krediteーションは大まかには完了し、次の段階として内部質保のオーディットを課題としていることもあれば、ドイツのようにオーディットはコスト面で意味のある一定規模以上の大学で重視され、一方、小規模の大学はむしろア krediteーションを残したがっているというような、並存の状態にあるところもある。

第二の論点は、学習成果の測定に関するものである。コンピテンシー (Competences) やラーニングアウトカム (Learning Outcomes) という概念で示されることがらについて、まず、そもそもそれらを測定することは本当に可能なのか、という問題がある。欧州各国では、そもそもそうした概念についてまだ共通理解を得るに至る状況にはなっておらず、理解の相違や誤解もある。また、大学を超えて、国を超えて共通の尺度を作成することの技術的制約も大きい。内部質保証の観点からは、それぞれの大学の責任において質を維持するということになりますが、しかし、はたして「知識、能力の涵養状況を測定して評価」することや「教育目標とそれに見合った学習成果の測定・評価」することが実現可能であるのか、仮に実現したとしても本来の教育意図をゆがめることになってしまわないだろうか、という指摘があつた。

第三の論点は、欧州と日本を比較して、日本の事情への示唆についてである。欧州各国では、学生や教員など多様な関係者を参加させることによって、制度の形骸化を一定程度防ぐことにつながっている。とりわけ、学生の参加ということが重要である。しかしながら、日本においては質保証について技術的な問題に関心が集中してしまう傾向があつて、どうしても、マネジメントのあり方や管理者の役割などの議論ばかりが前面に出てきている。多様なステークホルダー、特に当事者たる学生のそれについての議論が乏しいのではないか、という指摘が出された。

### (3) 米国における教育評価

次に、2011年7月15日の福留東土氏の報告「米国における教育評価」の要点と、それをふまえたディスカッションで出た論点をまとめる。

福留氏の報告の要点は次の通りである。現在の米国における教育の質保証の動向は、何よりも厳格な内部質保証、自己点検・評価である。アカウンタビリティの重視が強まる中で、米国において伝統的に実施されてきたア krediteーションに対して質保証の観点か

らの見直しの必要性が主張されてきた。それは、大学が持つ資源を対象としてきたことの意味を問い直し、大学の活動実態および成果へ焦点を移すべきという主張に表れている。学士課程教育における学習成果の重視という点もそこで主張されている。

米国のアクレディテーション団体の特徴を簡単に確認しておく。それは、「一様ではない」という点である。まず、評価主体の性格に違いがある。一つには、機関別、地域別の団体があり、これは高等教育や中等教育関係者によって政府には依存しないという意味で「自主的」に結成したものである。もう一つには、専門分野別の団体があり、これは専門職団体がイニシアティブをとるものが多数である。そして、こうした評価主体の性格の違いに合わせて、評価も対象とする範囲も大きく異なっている。そのことが、「一様ではない」という意味である。

また、アクレディテーション団体を統括・調整する団体がある。アクレディテーションは自主的に行われてきたという経緯があるので、非常に分権的で独自性が強いという性格がある。そこで、政府との関係という意味で、諸団体を統括・調整する団体が必要になる。これまで、いくつかの団体の設立・解散を経て、1996年 CHEA という団体が設立されて、各団体の認証を行うようになった。

そして、アクレディテーション団体を統括・調整する団体だけでなく、連邦教育省もまた各アクレディテーション団体を認証している。もともと連邦教育省は、もともと全米レベルの教育に関する情報やデータ収集や教育の側面支援が主な活動であった。つまり、直接の管理を行わない、という性格であった。ところが、1960年代くらいから、連邦政府の資金援助、たとえば研究資金や奨学金が各機関にとって重要性を増してきて、規制や監督を通じたアカウンタビリティが重要視されるようになってきた。そこで、連邦教育省による認証が行われるようになった。各機関は連邦教育省による認証を受けることによって、いろいろな資金援助プログラムへの参加が可能になるという仕組みである。これによって連邦政府がアクレディテーション団体や各大学の「自主性」に影響を及ぼすようになってきた。とりわけ 1990 年代以降、高等教育法の改正が行われて、認証を通じた統制の強化とも見られるような状況が生じてきている。

以上の福留氏の報告をもとにして、第1グループの参加者のディスカッションでは次のような論点が出された。

第一に、日本ではまだよく知られていない「アカデミックプログラムレビュー」をめぐって質疑がなされた。アカデミックプログラムレビューは内部質保証の重要な要素の一つで、とりわけいわゆる上位大学ではきわめて重要視されている。上位大学にとっては、アクレディテーションよりもむしろその水準を上回る要求を課すプログラムレビューが重視されている。しかし、日本においてはまだほとんど取り組まれていない。

ただし、そうした米国の動向の背景をふまえておく必要がある、それを可能にする文化があったことを見逃すわけにはいかない。たとえば、米国大学協会（AAU）には約 60 の研究大学が加盟していて、そこでは質に関する大学間の競争構造が根付いている。また、



連邦政府との関わりでいえば、戦後の GI ビルに基づいて政府が支出する資金について、アカウンタビリティを求められるようになった長い歴史があるという背景もある。

第二に、質保証という場合の「質」の定義に関わる議論がなされた。アクレディテーション団体は質の保証と質の改善とを目的とするわけだが、この両者の意味はまったく異なる。「質の保証」は、大学の外部への説明を目的としたものであって、最小限の基準を設定するものである。一方、「質の改善」は、大学の内部で行われるものであって、細かい基準を設定してなされる。この両者を区分けしたうえで、どのようなバランスを図るのかという問題を考えないといけない。

さらに、ラーニングアウトカムについてもその測定の目的は多様である。米国についていうと、連邦政府はアカウンタビリティを目的とするのに対して、アクレディテーション団体と大学は教育の質の改善を目的としている。先ほどのラーニングアウトカムが測定可能かどうかという論点とともに、その目的を明確にしておく必要があるという視点も、今後の評価のあり方をめぐる議論に活かされねばならない。

#### **（４）まとめ——内部質保証への着目**

以上（１）～（３）をふまえて、簡単なまとめと若干の問題提起をしておきたい。

欧米の動向を見るかぎり、全体として、内部質保証（Internal Quality Assurance）が鍵となる概念になっている。欧州は全体としてそうであるし、米国でも連邦政府は外部からの統制を強めようとしているが、アクレディテーション団体は機関内部での自律的な教育の改善向上への関心を強めているといえる。

近年の内部質保証への重点のシフトは、大学自身の内部からの自律的で自発的な動きこそが、「質」の保証の中心だというふうと考えられるようになってきている。内部質保証とは、たとえば、大場（2009）がユネスコ・ヨーロッパ高等教育センターをひいて紹介している定義では、「機関（プログラム）の一連の活動に関する質の監視（Monitoring）と向上（Improvement）に用いられる大学内部の仕組み」であるとされている。大学内部の仕組みをどう作るのかが重要な課題だが、たとえば、鳥居・山田（2010）は、教学 IR と FD の連動を梃子とした内部質保証のシステム化こそが、「学習・教授の質的な改善を第一義的な目的とした自己点検・評価活動を恒常的かつ磐石なものにする」ものである、と述べている。

以上をふまえて、最後に２点、若干の問題提起をしておく。

一つは、第１グループでの議論を通してくり返し出てきた論点は、いかなる仕組みであれ、評価システムの形骸化をどのように防ぐべきかという点を考えないといけない、ということである。これは、たとえば「監査の儀式化」の問題としてとらえ直すことができる。Power（1997＝2003）は「監査社会」と呼べるような変動が起きているけれども、監査一般の営みは儀式化してしまうことがある、と警鐘を鳴らしている。仮に監査体制が不十分であって、監査の失敗が認識されるようになったとしても、それは中止に向かうのではな

く強化が目指されることになってしまい、そのための内部統制が強まっていきがちである。その際、監査証拠の収集は莫大な作業量であって、時間と人員の大きなコストを払いながら、結局はデータの収集行為が目的化、儀式化してしまう危険性がある。また、必ずしも意図していなくてもそれが経営ツールとして利用されることになる。そのことの是非の問題もはらまれている。

Power はまた大学教育の質の監査にも言及している。「質の概念は、成果を強調する定義と成果を決定するプロセスを強調する定義との間で不安定にさまよっている」と述べている。まさに、大学評価の現状を的確に示しているようにも思われる。

もう一つは、「質」「評価」の概念のあいまいさの問題である。「質」や「評価」という概念を自明視してしまわないで、どのような評価システムの設計や運用においても、常にその概念の中身を問い直していくことが必要だ、ということである。羽田貴史（2004）は、「質」は歴史的・相対的な概念として捉えなければならず、それは単一の尺度によって規定されるものではないと述べている。すなわち、「質」とは、①学習や研究の成果のような伝統的な概念、②アカウンタビリティとしての基準・仕様への適合、③目的への適合性、などがありえるのあって、何を指すのかについてきちんと議論して共通の理解を得ることが必要なのである。

羽田（2009）はまた、「評価」の概念についても同様に慎重に考えないといけないことを指摘している。それは、Evaluation なのか、Assessment なのか、Accreditation なのか。また、大まかな傾向として、業績評価、数量化、格付け、外部評価へのバイアスがあるので、そうしたことへの慎重な目配りもまた必要なのである。

## おわりに

1995年に創刊された *Quality in Higher Education* 誌に掲載された15年間分の研究論文をレビューしたハーヴェイらによれば、掲載された諸論文を通して明らかになるのは、アカウンタビリティを強い要求するような外部評価は、全体として大学の質を高めることには成功してこなかったということである。それらはかえって、「信頼の解体」(dissolution of trust)を生んできた (Harvey and Williams 2010a)。また、高等教育機関が自らの手で改善向上していくという内部質保証も、特に消費者主義的なアプローチは大学人には受け入れられず、抵抗が続いてきた。一方では改革への熱狂やたくさんアイデアが出され、他方では旧来からの惰性 (inertia) や従順な無関心 (compliant indifference) が存続した15年間であったと結論づけられている (Harvey and Williams 2010b)。考えてみれば、日本の状況もよく似ている。

われわれは「何をすべきなのか」を議論する前に、「何をしたいのか」についてきちんと議論していく必要がある。あいまいな概念をスローガンのように語り、不明確な目的と未成熟な方法とで評価のシステムをいじり回していったとしても、決して望ましい未来は得

られない。欧米の動きから学ぶべきものがあるとすると、それは評価システムや評価手法などの持つ可能性と限界の両面でなければならない。欧米の経験をその両面で吟味しながら、日本の大学の現状の文脈を十分理解し、「われわれが何をしたいのか」という点に即した目的や手段の選択をすることが何よりも必要であろう。

第1部の以下の章では、本調査における第1グループの何人かに寄稿をお願いしている。第4節で簡単に触れた欧米の教育研究評価の現状については、研究会当日の報告をもとに福留東士（広島大学）・大場淳（広島大学）の両氏にあらためて論文のご寄稿をお願いした（第1部第2章・第3章）。また、第1グループでは、さらに日本における認証評価制度のあり方について議論を進め、日本学術会議で作成中の分野別参照基準との関連についても議論した。それらは、第1部第4章・第5章で論文としてまとめられている。また、検討の過程で近年注目を集めているアウトカム評価をどうみるかも重要な論点の一つとなった。これについては、本章でも触れているが、藤田英典氏（共栄大学）からも独立した論文をご寄稿いただいた（第1部第6章）。

## 注

- 1) なお、この議論に続けて、喜多村は、受験産業によるランキングなど「非政府機関による大学評価」は、「その評価は個々の大学のごく一面の部分的に偏った評価であるにすぎない」（喜多村 1999=2003, p.234）と述べている。評価尺度としてのランキングの持つ問題点については、本文で触れたとおり、他にも多くの論者が述べている。
- 2) ここでの金子の議論は、設置認可制度が戦後ある時期まで教育の質を保障する唯一の仕組みだったように論じているが、それはやや単純すぎる。むしろ金子自身が別稿で述べているように、アカデミックな水準を重視した教授会自治が、ある時期まではもう一方の質保証の役割を果たしたと考えるべきであろう（金子 1991=2003）。大学評価導入以前の教育がもつ難点は、質が低いのではなく、教授内容にせよ、学生の自主的学習への期待にせよ、質が高すぎて高等教育の大衆化の現実との間で距離が広がったと考えるべきだろう。現在でも、「ダメな授業」として描かれる事例のある部分は、質が低いのではなくて、むしろ講義の質が高すぎて学生の意欲や知識とのギャップが大きいものが含まれている。
- 3) 2006年のスプリングス報告以降の米国大学におけるアウトカム測定の現状を考察したKatz (2010) は、①技術的未熟さ、②費用、③組織内の不同意、④多様な組織の違いや役割を捨象してしまう問題、の4つを問題点として挙げ、本来は形成的評価こそが必要なのに、ジェネリック・スキルを測るテストはそれには使えないというふうに厳しく批判している。

## 引用・参考文献

- Clarke, Gill、2007、「イギリス高等教育における質保証」（吉川裕美子訳）『大学評価・学位研究』6、3-24。
- 江原武一、1994=2003、「大学評価の意味」大南正瑛編『文献選集 大学評価』エイデル研究所。
- 福留東土、2005、「米国のアクレディテーションにおけるアウトカム評価の動向」『COE 研究シリーズ 16 高等教育の質的保証に関する国際比較研究』広島大学高等教育研究開発センター、161-188。
- Harvey, Lee, and Williams, James, 2010a, “Fifteen Years of Quality in Higher Education,” *Quality in Higher Education*, 16(1): 3-36.
- Harvey, Lee, and Williams, James, 2010b, “Fifteen Years of Quality in Higher Education (Part Two),” *Quality in Higher Education*, 16(2): 81-113.
- 羽田積男、2012、「学生の学習成果評価と揺籃期 NILOA の機能」『教育学雑誌』（日本大学教育学会）47。
- 羽田貴史、2004、「大学組織の変容と質的保証に関する考察」『COE 研究シリーズ 8 高等教育システムにおけるガバナンスの組織の変容』広島大学高等教育研究開発センター、1-18。（再録：2009、塚原修一編『リーディングス日本の教育と社会 12 高等教育』日本図書センター、33-50。）
- 羽田貴史、2009、「質保証に関する状況と課題」羽田貴史・米澤彰純・杉本和弘編著『高等教育質保証の国際比較』東信堂、3-19。
- 広田照幸、2009、『ヒューマニティーズ 教育学』岩波書店。
- 星千枝・鈴木尚子、2009、「社会人に求められる能力の育成とアセスメント—イギリス・オーストラリア・アメリカの状況と日本への示唆」『BERD』16、48-56。
- 金子元久、1991=2003、「高等教育の構造変化と大学評価」大南正瑛編『文献選集 大学評価』エイデル研究所。
- 金子元久、2007a、「大学ランキングと大学」『IDE 現代の高等教育』495、4-8。
- 金子元久、2007b、「国立大学法人の評価—何が課題か」『IDE 現代の高等教育』490、4-12。
- 金子元久、2009a、「大学教育の質向上のメカニズム—『アウトカム志向』とその問題点」『大学評価研究』8、17-29。
- 金子元久、2009b、「基調講演 大学の教育力と学生調査」『一橋大学大学教育研究開発センター全学 FD シンポジウム報告書』10、5-34。
- 金子元久、2011、「質保証の新段階」『IDE 現代の高等教育』533、4-11。
- Katz, Stanley N., 2010, “Beyond Crude Measurement and Consumerism,” *Academe*, 96(5): 16-20.
- 川嶋太津夫、2009、「アウトカム重視の高等教育改革の国際的動向—「学士力」提案の意義と背景」『比較教育学研究』38、114-131。

- 喜多村和之、1999=2003、「大学評価の可能性—自己点検・評価と外部評価」大南正瑛編『文献選集 大学評価』エイデル研究所。
- 葛城浩一、2011、「ボーダーフリー大学教員の大学教授職に対する認識—『大学教授職の変容に関する国際調査』を用いた基礎的分析」『大学論集』42、159-175。
- 小林哲夫、2007、「偏差値で上位にない大学へ光を当てる」『IDE 現代の高等教育』495、25-30。
- Newton, Jethro, 2002, “Views From Below: Academics Coping With Quality,” *Quality in Higher Education*, 8(1): 39-61.
- Newton, Jethro, 2010, “A Tale of Two ‘Quality’s’: Reflections on the Quality Revolution in Higher Education,” *Quality in Higher Education*, 16(1): 51-53.
- 日本学術会議、2010、『回答 大学教育の分野別質保証の在り方について』日本学術会議大学教育の分野別質保証の在り方検討委員会。
- 大場淳、2009、「フランスにおける高等教育の質保証」羽田貴史他編著『高等教育質保証の国際比較』東信堂、177-195。
- Power, Michael, 1997, *The audit society: rituals of verification*, Oxford: Oxford University Press. (=国部克彦・堀口真司訳、2003、『監査社会—検証の儀式化』東洋経済新報社。)
- 坂本達也、2007、「慶應義塾の国際戦略と大学ランキング」『IDE 現代の高等教育』495、47-51。
- Stensaker, Bjorn, 2007, “Quality as Fashion: Exploring the Transformation of a Management Idea into Higher Education,” Don F. Westerheijden, Bjorn Stensaker, Maria Joao Rosa, (eds.), *Quality Assurance in Higher Education: Trends in Regulation, Translation and Transformation*, Dordrecht: Springer-Verlag.
- Taylor, Paul and Braddock, Richard, 2007, “International University Ranking Systems and the Idea of University Excellence,” *Journal of Higher Education Policy & Management*, 29(3): 245-260.
- 鳥居朋子・山田剛史、2010、「内部質保証システム構築に向けた教学 IR と FD の連動」『大学教育学会誌』32(2)、39-42。
- 山田礼子、2009、「学生の情緒的側面の充実と教育成果—CSS と JCSS 結果分析から」『大学論集』40、181-198。
- 山田礼子、2011、「大規模継続学生調査の可能性と課題」『大学論集』42、245-263。

## 第2章 米国のアクレディテーション——日本の認証評価との比較の視点から——

福留 東土（広島大学）

### はじめに

本稿では、アメリカの大学評価制度について、日本との違いを主な観点として考察を行う。アメリカにおいて日本の認証評価に当たるものは、アクレディテーション（*accreditation*）になるが、これについて中心的に取り上げる。ただし、後半では大学の機関内部の質保証に関する事例を取り上げてみたい。今後の日本の認証評価、質保証のあり方を考えていく上で非常に重要な事例ではないかと考えるからである。

#### （1）米国高等教育質保証の動向と枠組み

本論に先立って、アメリカの高等教育質保証あるいは評価をめぐる最近の動向をおさえておきたい。

第一に、1990年代以来、高等教育機関に対するアカウンタビリティが一貫して強調される傾向がみられる。第二に、高等教育の質保証の手段としてのアクレディテーションのあり方に対する批判が非常に強まっている。特に、2006年に出されたアメリカ連邦教育省高等教育将来構想委員会による報告書“*A Test of Leadership*”、いわゆるスプリングスコミッションのレポートは、アクレディテーションや高等教育の質のあり方に対する厳しい批判を含むものであり、日本でもいろいろなところで紹介されてきた。

批判のひとつとして、アクレディテーションは、大学の内部の観点に非常に偏っていて、社会や消費者の観点に十分に応えきれていないという見方がある。もうひとつは、これまでのアクレディテーションはインプットベースで、大学がどういう資源を持っているかということに主な重点を置いて評価が行われてきたことに対して批判が出ている。そうした指標では大学の活動の実態がなかなか見えないし、その活動を通してどういう成果を上げているのかが十分に明らかになっていないといった批判である。

第三に、高等教育のなかで学士課程教育が大きな焦点とされるようになってきている。時に、上記の動向とも関係して、学生の学習成果が重視されるようになってきている。

こういった大きな流れは、それぞれの大学に対して厳格な内部質保証、あるいは自己点検・評価を要求する動きともつながっており、マクロに見れば、日本の文脈ともかなり共通する部分がある。

しかしながら、日本とアメリカで根本的に違う点がある。それは、政府が高等教育の質保証にどう関与するか、という点である。

アメリカでは、政府が規制を強めていこうとするこれまでの動向に対して、アクレディテーションに関わる団体の側は、政府による規制に強く反発を示している。ただし、その一方で、政府の規制を排除するために、アクレディテーションなり、高等教育機関、高等教育の内部者によって、自分たちの手できちんと質保証ができることを示していく動きが進んでいる。機関レベルの質保証が重要になってきているという文脈は、大学団体や高等教育機関の側の自律的で対抗的な動きであると読むことができる。

図1は、米国高等教育の質保証の基本的な枠組みを示したものである。

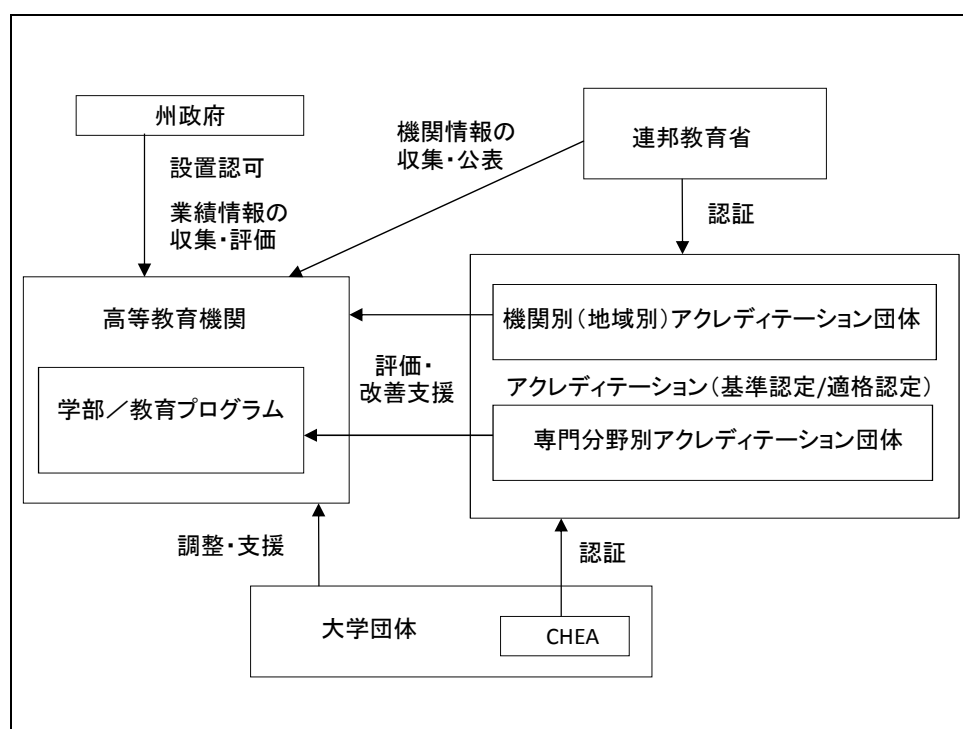


図1 米国高等教育の質保証の基本枠組み

高等教育の質保証、あるいは高等教育のガバナンスの基本的な枠組みを大きくまとめると図1のようになる。機関レベルのアクレディテーションでは、6つの地域別に分かれたアクレディテーション団体がある。また、専門分野別のアクレディテーションは、対象となる分野が多く、団体の数も非常に多くある。こういう2種類のアクレディテーション団体があって、前者が機関全体を、後者が個別の学部や教育プログラムを評価対象にしている。

アクレディテーションに対しては、連邦教育省が認証活動を行っている。また、大学団体のひとつとして、アクレディテーション協議会（Council for Higher Education Accreditation, CHEA）という団体があり、これは政府とは別の形でアクレディテーション団体の認証を行っている。

また、質保証とはやや異なるが、アメリカには大学関連団体が非常に多くあり、全米レベルの大学ガバナンスにおいて、それらが非常に重要な役割を果たしている。連邦政府が

高等教育の規制をすることは、直接的にはない。もちろん、州は直接管理を行っているけれども、連邦レベルでは間接的な関与が原則である。しかし、後述するように、連邦政府の関与と影響力は次第に強まってきている。

## **（２）アクレディテーションの基本的特質**

米国のアクレディテーションの基本的な特質は、よく言われる通り、ボランティアにある。アクレディテーションの持つ基本的な性格として、「大学関係者が自ら結成したボランティアな団体によって質保証が行われる」、「非政府の質保証である」、「機関による自己評価と同僚のピアレビューによって行われる評価である」、「自己規制あるいは改善を志向している」といった点が日本でも紹介されてきた。

その一方で、特に1990年代以降は、連邦教育省が、先ほど述べた、アクレディテーションに対する認証をひとつの媒介にして規制を強めてきている。連邦政府から認証されたアクレディテーション団体が、各教育機関を基準認定する。高等教育機関はその基準認定（アクレディテーション）を通して連邦政府の資金援助プログラムへの参加資格を得る、という形になっている。それゆえ、かりにそのアクレディテーション団体が政府の認証を失ってしまうと、大学が連邦政府からの資金の受給資格を失ってしまうことになる。そこが、アクレディテーションにとっても非常に重要な側面になってきている。特に、1980年代後半頃からアカウントビリティの要求が非常に強まっている。

それゆえ、ボランティアはアクレディテーションの基本的な精神として踏まえておくべきだが、現実には、政府との一定の関係のなかでアクレディテーションが質保証の役割を果たしているといえる。アクレディテーションについては、「パブリック・プライベート（public-private）」という言われ方をすることがある。もともとはプライベートだが、パブリックの側面もかなり強くなってきているということである。ここでのパブリックには、おそらく「公共」という意味と「政府的な」という意味の両方が含まれている。

## **（３）本論文のねらいと構成**

以上のことを前提にしながら、本稿では、次の5つの問題を考えてみたい。第一に、より具体的に、どのようなプロセスの中で、アクレディテーションの特質が形成され、あるいは変化してきたのか。第二に、アクレディテーションと政府の関係をどうとらえるか。第三に、アクレディテーションに上記の通り大きく二つの形態があるとする、その形態ごとに何か特質があるのかどうか。第四に、アクレディテーションの基本的な理念である「ボランティア」とはどういうものか。それらの検討を通して、最後に、大学が自治に立脚しつつ、大学の公共性を担保できるような評価システムはどのように可能なのか。私自身は、評価や質保証の基本的な目的は、それらを通して大学が自治を行い、自律的に自分たちの質を保証していくことによって、大学が公共的な存在であることをきちんと示すことではないかと考えている。それを問題意識に置いているので、最後にその点を少し考



えてみたい。

本稿の構成は次の通りである。まず、アクレディテーション団体の特徴を簡単に検討する（第1節）。次に、政府との関係について考察を行う（第2節）。さらに、アクレディテーションの統括や調整の問題を考える（第3節）。そして、アクレディテーションとは異なるけれども、大学の内部質保証について、質保証あるいは評価の一環として事例を紹介する（第4節）。

## 第1節 アクレディテーション団体の特徴

### （1）機関別（地域別）アクレディテーション団体

米国のアクレディテーション団体をおおまかに分類すると、機関別の評価を行う地域別のアクレディテーション団体と、専門分野別アクレディテーション団体の2種類がある。細かく言うと、もう一種類、地域別ではなくて全国レベルで機関単位のアクレディテーションを行う団体もある。宗派立の機関を対象にする団体や、職業教育機関の評価をやっている団体などが、そういう形態を採っている。しかし、議論を整理するため、ここでは主要な形態として大きく上記の二種類に分けることにする。

まず、機関別・地域別のアクレディテーション団体について説明する。機関別・地域別のアクレディテーション団体は、地域ごとに若干ばらつきがあるけれども、19世紀の後半から20世紀の初頭にかけて作られている。その際に、団体の結成に関与したのが大学の関係者と中等教育の関係者であった。当時、アメリカでは高等教育が大きく拡大していた。学校制度自体が19世紀の後半からかなり確立してきたのだが、それまでの高等教育は、就学年齢もかなりばらばらで、各大学の入学要件にもかなりのばらつきがみられた。そういったことに対して、制度的に次第に発達していた中等教育の関係者からの不満があった。また、中等教育機関の教育のあり方にもばらつきがあったので、高等教育の側も、各中等教育機関がどういう教育をしているか、その実態がなかなか掴みにくかった。そこで、中等教育とは何であるのか、高等教育とは何であるのか、その接続関係をどう整理するのかという問題が持ち上がってきた。

ここでも、米国固有の歴史的な背景があり、政府に特に依存せずに、教育関係者が自主的にそういう動きを見せたという意味では、ボランティアな活動であったといえる。ただし、単純に教育の質を高めていくこと以上に、今述べたような中等教育と高等教育との接続関係の整理という明確な必要性があったことは強調しておく必要があるだろう。

1890年代から第一次世界大戦あたりまでの20数年の時期は、歴史家のロジャー・ガイガーによれば、高等教育の成長と標準化の時期であると位置づけられている。高等教育システムが、一方で急激に拡大しつつ、他方でそれをどう整理してシステムとして標準的なものにしていくかという課題を抱えていた時期だということである。そこで、この時期に大学団体が数多く結成され、また、専門分野の学会、さまざまな学術団体もこの時期から

活動を始めていった。

そうした流れのなかで、ア krediyteeshon も、中等教育、高等教育、及びその接続関係について、それをいかに基準化していくか、あるいは、それぞれの教育をいかに制度化していくかという必要のなかから出てきたということができる。

## **(2) 専門分野別ア krediyteeshon 団体**

もうひとつの専門分野別ア krediyteeshon 団体は、これまで日本ではそれほど紹介されてきていない。しかし、その歴史を見ていくと、専門分野別のア krediyteeshon 団体の存在は、いろいろな複雑な問題をはらんでいる側面があると考えられる。

その特徴は、まず、ほぼ「専門職」分野に対応している点である。それぞれの分野によって評価対象となる範囲は多少異なるけれども、基本的にはプロフェッショナルスクール、およびその内部のデパートメントや教育プログラムが対象となる。つまり、研究大学でいうと、大学の最も中心になる組織としてカレッジ・オブ・アーツ・アンド・サイエンスとか、スクール・オブ・アーツ・アンド・サイエンスがあるけれども、それらは基本的には専門分野別ア krediyteeshon の対象にはならない。

専門職のア krediyteeshon 団体がどういうかたちで設立されたかということについては、団体によっていろいろなタイプがある。最も多いのは、専門職の側がイニシアティブを執ってア krediyteeshon 団体を作っていくというタイプである。もちろん、大学関係者が中心になって作られたものもあるけれども、特に、医療系の分野や、法学など、専門職の資格との結び付きが非常に強いのが一般的である。

すなわち、専門分野別ア krediyteeshon は、必ずしも大学関係者の意向によるだけではなくて、専門職側の要求やその専門職の論理がかなり反映されている。特に、新しい専門職にとっては、ア krediyteeshon という制度が専門職としての自身の存在を鮮明にしていく手段にもなりうる。それゆえ、社会が高度化していろいろな専門職が新たに登場してくるなかで、ア krediyteeshon 団体がどんどん増えていくことになり、大学は、数多くのア krediyteeshon に対応していかなければいけなくなっている。

専門分野別のア krediyteeshon は——専門職大学院を対象とする日本の認証評価でも基本的に同じだと思われるが——、部局レベルあるいは教育プログラムレベルの評価に関して、評価対象の教育研究活動の中身にかかなり踏み込んで評価を行うのが一般的である。そのため、たとえば、プロフェッショナルスクールのディーンなどに聞くと、機関単位の評価に比べると評価としての有効性が高いという回答が見られる。

もうひとつの現実的な機能として、部局ごとの専門分野別ア krediyteeshon を通して、全学に対して部局の現状をア krediyteeshon にいわば代弁させるといえる。それを全学レベルの執行部、大学経営者に対して主張していく機能も持っているといえる。

以上のように、ア krediyteeshon 団体といっても、機関別・地域別のア krediyteeshon 団体と、専門分野別ア krediyteeshon 団体の2種類の団体の性格は、かなり

異なっている。単純に評価対象の範囲だけの問題ではなくて、その団体の性格にも、かなり大きな違いが見られるわけである<sup>1)</sup>。

## 第2節 政府との関係

次に、政府との関係について考察する。まず、高等教育に対する政府の役割に関していえば、先述した通り、連邦政府は直接の管理・規制は行わないのが原則であり、高等教育機関の設置認可は州政府が行っている。しかし、州の設置認可は非常に緩やかなかたちで行われることが多いとされている。

連邦教育省の役割は、元来は全米レベルの教育に関する情報やデータ収集による現状把握、あるいは教育に対する側面支援が中心であった。ところが、1930年代頃から連邦政府は高等教育への資金提供を行うようになり、1960年代からは、特に研究資金や学生に対する奨学金の提供が大規模化されていった。そういう資金援助が重要性を増すにつれて、連邦政府による規制や監督を通じたアカウントビリティが重要視されるようになってきた。

連邦教育省によるア krediteーション団体の認証が意味しているのは、学生奨学金など、連邦の多様な資金援助プログラムへの参加を担保するということである。ア krediteーション団体が連邦教育省に認証されるということが、ア krediteーションを受けることによる資金面での具体的メリットとの結び付きを生じさせることになった。そのことが、本来ア krediteーションが持っているボランタリズムに、かなり大きな影響を及ぼす結果になっている。

とはいえ、政府との関係については、日本の状況とはかなり違うことを感じさせる点もある。たとえば、専門分野別のア krediteーション団体のうち、いくつかの主要な団体で、連邦教育省による認証を受けていない場合がある。一時期は受けていたけれども現在は受けていない団体もあるし、一貫して連邦教育省の認証を受けていない団体もある。たとえば、AACSBは経営学のビジネススクールの基準認定を行う団体だが、現在、教育省による認証を受けていない。ABETは工学分野の団体で、日本のJABEEに対応するものだが、それも同様である。

こういった主要な団体は、評価対象となる機関全体がすでに機関単位でのア krediteーションを受けていることを自分たちのア krediteーションを受けることの前提条件にしている。だから、その条件を満たしている時点で既に連邦プログラムへの参加は担保されている。それゆえ、自らが連邦教育省の認証を受けなくても、具体的なデメリットは生じないわけである。

また、ア krediteーションと政府との関係で特筆すべきは、日本でも既にいくつかの論文で紹介されているが、1990年代以降、高等教育法の改正が何度か進んできて、そのなかで、認証活動を通じた規制強化の動きが見られるということである。

政府からの資金援助という具体的なメリットとア krediteーションとの結び付きをど

う見たらよいのかという点は、私自身まだ十分に明確な見解を持っていない。政府が一方的にアクレディテーションを利用しているという見解もある一方で、アクレディテーション団体や大学にとっても、政府の認証を通して社会の信認を得るという意味合いがあるという見方も可能である。おそらく、その両方の側面があるのではないかと思われる。

### 第3節 アクレディテーションの統括・調整

次の論点は、アクレディテーションの統括・調整をどう見るか、という点である。米国のアクレディテーションの仕組みは、先述のように基本的にボランタリーなものなので、非常に分権的であり、個別の団体が強い独自性ないしは個別性を持っている。それゆえ、全米レベルでの統括や調整などにはあまりなじまないのが本来の性格であろう。

しかし、近年はいろいろな必要性からそういう動きが出てきている。しかし、そこではやはり非常に難しい問題が生じているようである。それを通して、米国のアクレディテーションの実態を見ていきたい。

まず、現在は、CHEA（Council for Higher Education Accreditation）という団体が全米のアクレディテーション団体の認証を行っている（図1参照）。政府が行う認証は、連邦の資金提供プログラムへの参加という具体的なメリットと結び付いているけれども、CHEAが行う認証は具体的なメリットはない。むしろアクレディテーションの世界での市民権を得るというか、きちんとしたアクレディテーション団体であることが認められる、といった意味合いがある。それゆえ、連邦政府によるアクレディテーション団体の認証とCHEAによる認証とは、認証の意味合いが異なっている。

地域別のアクレディテーション団体でヒアリングをすると、やはり連邦政府の認証を受けているかどうか非常に重視されていることがわかる。それが、自分たちが大学に与える具体的なメリットとつながっている、というのである。逆に、専門分野別の団体にとっては、政府の認証は重要ではなくて、むしろ、CHEAによって認証されて、市民権が得られていればそれでよいというふうに回答されることがある。

とはいえ、専門分野別アクレディテーション団体には実に多種多様な団体が存在しており、社会における自分たちの存在を明確にするために、どちらからでもよいからとにかく認証を取ることを目指すといった団体もある。

現在のCHEAにみられるようなアクレディテーションの統括あるいは調整は、基本的には第二次世界大戦後になって登場してきた。

その背景として第一に、専門分野別アクレディテーション団体が非常に増えてきたので、それを抑制するという必要性が浮上したからである。個別の専門分野別アクレディテーションを受けることは、大学に対して資金的あるいは時間的なコストが非常に掛かる。また、より大きな問題として、外部の専門職が要求する基準によって大学のカリキュラムを組んだり、資金的にいろいろな手当てをしたり、施設・設備を充てたり、教職員を雇わなければ

ばいけなかったりという問題があった。大学側からは、これが大学の自律性に対する不満として存在していたようである。

第二に、地域別の団体では、それぞれの団体がある程度独自に評価を行っていたけれども、それを全米レベルである程度統一化、あるいは標準化していく必要性が出てきたという背景もあった。特に、政府との関係のなかで、ア krediteーションの存在意義を明確にしていくために、緩やかながらもある程度共通の基準で質を担保していることをきちんと示す、という必要性も高まっていた。

ア krediteーションの統括・調整に関わって、1949年に二つの団体ができた。ひとつ目の団体はNCA（National Commission on Accreditation: 1949～1975）である。これは大学の学長団体として創設され、ア krediteーション団体は、この団体には入っていなかった。この団体は、先に触れたような専門分野別の団体の増殖に対して、むしろ高等教育機関の側が主導権を握ってア krediteーションのシステムを作りたいと考えていた。メンバーとして、高等教育機関とア krediteーション団体以外のさまざまな大学団体が入った。

もうひとつは、FRACHE（Federation of Regional Commission of Higher Education: 1965～1975）、前身はNCRAA（National Committee of Regional Accrediting Agencies: 1949～）である。FRACHEは地域別ア krediteーション団体の連合体である。これは、先述したように地域別ア krediteーションを全米レベルである程度統一化して標準化していく必要性に応じて出てきた団体であり、各団体の自主性を尊重しつつ、調整によって全米的な標準化を狙っていた。しかし、この時点では、そういった調整は十分に機能しなかったようである。

1975年になって、それらふたつの団体が今度はCOPA（Council on Postsecondary Accreditation: 1975～1993）という、単一の大きな団体に統合された。これは、名称に「ポストセカンダリー・ア krediteーション」と書いてあるように、大学以外の職業学校や営利学校といったポストセカンダリーの教育も含んだ、非常に包括性の高い団体であった。

この団体の成立には、それ以前の時期の反省を踏まえて、もっと一元的な組織を作らなければならないという意図があり、ここにはア krediteーション団体も入るし、学長団体も会員として入ることになった。ところが、多種多様なメンバーを含むので、逆に複雑な利害の対立がひとつの組織のなかに内包されてしまうことにもなった。

たとえば、奨学金の返還の滞納問題がある。特に、営利学校や職業学校では、奨学金の返還滞納が非常に多くあった。ポストセカンダリーということで、そういった学校も傘下に含んだのであるが、そういった種類の学校が問題を抱えてしまうと、それがア krediteーションの世界全体に大きな影響を及ぼしてしまうことになっていった。

そのような利害対立によって、地域別の団体がCOPAを脱会し、結局COPAは解散するという経緯をたどった。その後も、多様な主体が関与する非常に複雑な議論の末に、

1996年になって今のCHEAが作られた。

こうした動きをどのように見たらよいのか。ア krediyteeshon が政府など外部との関係をどう持つかを考えていくと、ア krediyteeshon 全体を統括したり調整したりする団体の存在は必要不可欠であろう。しかし同時に、ア krediyteeshon の自主性や独自性をいかに推進するかということと、それを全体としてどう統括・調整していくのかということの二面性の難しさがあるといえる。そのような葛藤とそのバランスを取ろうとする過程として、ア krediyteeshon の発展のプロセスを見ていくことができるだろう<sup>2)</sup>。

もうひとつ、論点として考えられるのは、大学にとってのア krediyteeshon のマネージの問題である。制度化されて発展してきたア krediyteeshon は、同時に、複雑化し、あるいは官僚化して硬直的なものになってきた側面もある。そのなかで、大学から見たときに、ア krediyteeshon をどうマネージしていくかが、非常に難しい問題として存在している。

## 第4節 内部質保証の事例

ここからは、大学における内部質保証の事例をひとつ紹介してみたい。ア krediyteeshon 団体のような機関外部からの質保証と関わって、それぞれの大学が内部質保証のあり方として、大学が自分たちの質を自分たちで定める主体的な姿勢が非常に重要だと考えるからである。

ここで紹介するのは、研究大学における自律的な内部質保証の事例であり、2011年3月にカリフォルニア大学バークレー校に行って調査をした「アカデミック・プログラムレビュー」と呼ばれるものである。評価対象となる教育研究単位は、基本的にはデパートメントである。プロフェッショナルスクールの場合は、スクール全体が評価対象となる。これらの単位ごとに、8年から9年おきにレビューが実施されることになっている。バークレーの担当者によれば、これは必ずしもア krediyteeshon の要求に応えるためにやっている活動ではなく、もともと1970年代から実施してきたものだそうである。自分たちは、ア krediyteeshon で要求されている水準以上のことを行っていると述べており、自律的な内部の仕組みが結果的にア krediyteeshon の要件に応える形になっているといえる<sup>3)</sup>。

評価内容については、それぞれのデパートメントがどういう使命を持っているか、どういう計画を持っているか、教育、研究、教員人事の問題、デパートメントの管理運営、職員や設備の配置はどうかといった、当該評価期間内の活動全般を対象にしている。

評価のプロセスは、細かく見ていくといろいろあるけれども、おおまかにまとめると次のようになる。まず、デパートメントが自己点検・評価を作成する。それに対して、大学外の研究者である外部評価者を中心に外部評価委員会を構成して、そこがサイトビジット

を含めて評価して、評価書を作成する。そして、レビュー参加者の間での協議などを通じて評価結果を確定し、どういう改善点が求められるかをデパートメントに報告する。そしてそれをフォローアップしていくところまでが、ひとつのサイクルとして見なされている。バークレーのアカデミック・プログラムレビューの特徴は、非常に多様な部署が関与しているという点にある。

まず、それぞれのデパートメントのレビューごとにプログラムレビュー管理委員会が作られ、そのなかに多様な委員が入る。まず、執行部のメンバーと評議会（Academic Senate）の代表が入る。評議会は、いわばバークレーの教員を代表する形で管理委員会に入っている。教員の代表として評議会のメンバーが入ることが重要な特徴といえるだろう。また、外部評価者が非常に重要な役割を果たしている。基本的には、バークレーと同等の水準の主要な研究大学から、同じ分野の教員3人から5人ぐらいが外部評価者として加わっている。

もうひとつの関連する部署としては、計画・分析室（Office of Planning and Analysis; OPA）がある。日本で最近、非常に注目されるようになったIR（Institutional Research）の機能を持った部署である。ここが、プログラムレビューのためのさまざまなデータや評価情報を当該のデパートメントや評価者に提供して、レビューを側面から支援していく形になっている。以上のように、外部の研究大学を含めた多様な部署のメンバーが関与する点に大きな特徴がある。

このプログラムレビューの特質と意義はどのようなものか。ひとつの分かりやすい意義としては、アクレディテーションが対象にしない部分を扱っている、ということである。先に述べたように、アクレディテーションは、機関単位と職業分野のプログラムレベルの評価しか行わない。専門職分野以外のプログラム単位でのレビューは、アクレディテーションでは基本的には行われないので、それが機関内部のこうしたプログラムレビューによって担保されているといえる。それが非常に実質的な評価になるという意味で、こうしたレビューの存在意義があると思われる。

もうひとつの特質としては、ピアレビューを基本としているということである。同じ分野の研究者が評価を行っているのである。

もうひとつの特質は、内部の教員が関与しているという点である。まったく遠い分野ではなくて、基本的には類似の分野を手がけ、当該の分野の内情が比較的理解できる教員が、バークレーの内部者として関与する。これは、バークレーという機関として各デパートメントに求める水準を外部者だけに任せるのではなくて、内部者もそれをきちんと担保することと、内部の人たちが関与することで、外部から見た場合には分かりにくい学内の文脈を外部評価者にきちんと伝えていく役割もある、とのことである。

もうひとつは、若干文脈の違う話になるが、先ほど言った計画・分析室（OPA）との関係である。ここでは、IR組織によるデータ提供を通じた現状把握が行われている。これは、評価プロセスの支援と、多くのデパートメントがあって、評価の数自体は大変な数

があるので、それを標準化し、効率化していく意味がある。プログラムレビューでは、レビュー項目として、ピアとなる研究大学との比較を通した検討が求められている。これは、他の大学の同じ分野のデパートメントと、バークレー内部の同じ系列のデパートメントという二つの比較を実施することになっている。その際、比較可能なデータをOPAが提供しているのである。

IRに関しては、現在の日本で研究が盛んに進んでいるが、そこでは全学の戦略にIRがどう支援するかという観点が非常に強調されているように思われる。もちろんそれがIRの最大の機能であると考えられるが、同時に、部局の活動とIRの担当部署がどうつながっているか、IRが部局の活動をどう支援できるかという点から考えていくうえで、バークレーの事例は参考になる。

また、プログラムレビューの際にどういう観点が強調されるかについて触れておく。もちろん、これまでの活動についてレビューを行うのだが、しかし、過去の活動についていろいろ細かく詮索するのではなくて、将来に向けた計画や、デパートメントがどういう方向に向かうべきかを念頭に置いて——forward-lookingという言葉が頻出する——、将来をどうするか、そのために過去や現在をどう評価するか、ということが基本的な観点とされている。

この事例がアメリカの大学の内部質保証の活動としてどのくらい一般化できるものなのかについては、まだ確かなことはいえない。しかし、内部質保証のあり方としては、研究大学としての自分たちの質を自分たちで定める主体的な姿勢が見られることが非常に重要で、これからの日本の認証評価を考えていくうえでも示唆があるのではないだろうか。今後は、この点についての各大学の活動に関する考察をより深めると同時に、従来日本ではあまり紹介されてこなかったこのような活動が、他の機関でどのくらい一般化されているのかについても、さらに調査していく必要がある。

## 注

- 1) 基本的な点として、米国のアクレディテーションは、法制化されているものではないので、受ける／受けないは大学の判断であるが、実質的には、受けなければ大学として生存していくことは難しい。ただし、地域別のアクレディテーションに対して、受けたくても受けられない大学はあるようである。申請はしても前段階で排除されてしまう大学がかなりあるとのことである。それゆえ、そういう大学は、高等教育機関として生存していくことはかなり難しいし、かりに生存したとしても、財政をきちんと保証していただけるだけの学生を引き付けることは難しい。地域別のアクレディテーションに関しては、法制化はされていないものの、実質的には、かなり義務に近い存在になっているといえる。ただし、架空のアクレディテーション団体にアクレディットされていると言っているディグリーミル、アクレディテーションミルも氾濫している。
- 2) スペリングスコミッションの報告書をめぐっては、CHEAがアクレディテーションを



代表するかたちで、かなり何度も反論をしたり、コメントを出したりするプロセスがみられた。CHEAや大学団体は、大学に政府からの要求をそのままストレートに流すのではなくて、たくさんのレポートを出すなどして、ア kredィテーション団体と大学の利益になるようなかたちに転換するべく活動を展開する役割を果たしたといえる。社会の要求・圧力に対して、一方でそれに対抗しつつ、他方でそれらを高等教育の実質的な改革にうまく利用していくという観点からみたとき、これら大学団体の活動の重要性を窺い知ることができる。

- 3) 米国の研究大学が加盟するAAU (Association of American Universities) という大学団体があり、62の研究大学が加盟している。厳格なメンバーシップで、加盟校数がほとんど増えないが、加盟大学間では、教育研究の質に対する競争構造がかなり存在している。つまり、自分たちの手で教育研究の質をきちんと伸ばしていかなないと、主要な研究大学として次第に取り残されていってしまうという、機関間の競争構造が各大学の質に対する努力を促している面がある。

#### 主要参考文献

- Alstete, Jeffrey W., 2007, *College Accreditation: Managing Internal Revitalization and Public Respect*, New York: Palgrave.
- El-Khawas, Elaine, 2001, *Accreditation in the USA: Origins, Developments and Future Prospects*, Paris: Institute for Educational Planning, UNESCO.
- Geiger, Roger L., 2005, "The Ten Generations of American Higher Education," P. G. Altbach et al. (eds.), *American Higher Education in the Twenty-First Century* (2<sup>nd</sup> Edition), 38-70.
- Orlans, Harold, 1975, *Private Accreditation and Public Eligibility*, Lexington: Lexington Books.
- Semrow, Joseph J., et al., 1992, *In Search of Quality: The Development, Status and Forecast of Standards in Postsecondary Accreditation*, New York: Peter Lang Publishing.
- University of California, Berkeley, 2010, *Guide for the Review of Existing Instructional Programs*.
- University of California, Berkeley, 2011, Academic Program Review Website, (<http://vpapf.chance.berkeley.edu/apr/>)
- University of California, Berkeley, 2011, Office of Planning & Analysis Website, (<http://opa.berkeley.edu/>)
- Young, Kenneth E., et al., 1983, *Understanding Accreditation*, San Francisco: Jossey-Bass.

### 第3章 欧州における大学評価の現状と課題——日本の認証評価制度との比較を視野に入れて——

大場 淳（広島大学）

#### はじめに

現在、欧州では、評価を含みつつ質保証を通して大学教育研究（特に教育）の在り方が議論されている。そこで、私の報告では質保証を主題として取り上げつつも、評価に重点を置いて検討していきたい。

#### （1）質保証制度の概観

福祉国家後の国家統治形態のあり方として、ニーブ（Guy Neave）は「評価国家」の概念を提示したが、その下で新公共経営（NPM）に基づく規制緩和や機関の自律性の拡大といった改革が行われ、事後的評価制度の整備が図られることになった。同時に、そうした改革の限界や改革によって引き起こされた弊害が指摘されることになる。市場化に伴う不適切な行動の顕在化、平等性や公平性に対する影響、新公共経営（new public management: NPM）に伴う主人・代理人のディレンマなどである。

そして「評価国家」後の欧州、特に、英国を除く大陸欧州のあり方について、二つのことを指摘しておきたい。第一に、欧州統合と地方分権の同時進行である。上から、あるいは、下からといった脱政府化が進んで、各国中央政府の権限が後退していく。それに伴って、大学運営に参加する利害関係者がさらに拡大し、中央政府や地方政府（ドイツのような分権制の場合）は数多くある利害関係者の一つになるのである。欧州は伝統的に中央政府の権限が強力であったが、次第に、それは後退していく流れになってきた。第二に、ボローニャ・プロセスとリスボン戦略といった全欧的取組である。2000年前後に、ほぼ同時に、欧州高等教育圏及び欧州研究圏の創設へ向けた政策が始められることになった。前者にかかるボローニャ・プロセスの重要な二つの要素が、学位構造の収れんと質保証にあることはよく知られていることである。これは、大学団体や学生団体等を取り込みつつ、欧州諸国の高等教育担当大臣会合による合意に基づく取組である。他方、リスボン戦略は政府間組織である欧州連合（EU）の行う政策である。人的資本論を背景として、大学は世界で最も競争力がある知識経済へ向けて貢献するといった政策がとられている。そのために、高等教育の修了者の拡大や、この戦略の目的に合致するような教育内容の提供を図ることが目指され、教育の質保証の中心概念の一つとして資質能力（competences）や学習成果（learning outcomes）に重点を置いた評価制度が構想されることになった。ボローニャ・プロセスにおける学位構造と密接にかかわる資格認定枠組み（qualifications

framework)でも同様に、資質能力や学習成果が重視されていくことになったのである。

質保証については、まず全欧的取組に触れておきたい。欧州高等教育質保証協会 (European Association for Quality Assurance in Higher Education: ENQA) の創設、欧州規準・指針 (European Standards and Guidelines: ESG) の制定、欧州高等教育質保証登録機構 (European Quality Assurance Register in Higher Education: EQAR) の設立、さらには、欧州大学協会のような既存の機関によって、さまざまな質保証のための活動が行われている。それと同時に、各国政府においても質保証機関の設置やその他の政策がとられている。たとえば、英国の質保証機関 (QAA) やオランダと蘭語圏ベルギーを合わせたオランダ＝フランデレン適格認定機関 (NVAO) が設置された。また、フランスでは、従前から存在していた大学評価委員会 (CNE) と政府のなかにあった教育課程 (プログラム) の認証を行う組織に加えて研究評価委員会を統合して総合的な質保証機関である研究・高等教育評価機関 (AERES) が設置された。フランスは非常に中央集権的な国であるが、AERES の評価結果はある程度予算と連動されている。従前の CNE が形式的な評価を行っていたのに対して、AERES は教員評価制度のメタ評価を含んで総括的な評価に重点を置くことになった。

## (2) 資質能力 (competencies) と学習成果 (learning outcomes)

学習成果と資質能力については、それぞれの要素を同定する作業が行われている。教育構造同調事業 (Tuning Project) や ダブリン記述書 (Dublin Descriptors) がそれに該当する活動であって、領域ごとあるいは階層ごとに同定することを試みている。それらを取りまとめたのが欧州レベルの欧州高等教育圏資格認定枠組み (Framework for Qualifications of the European Higher Education Area: FQ-EHEA) で、これを基に各国において国内枠組みを作っていくことが目指されている。

ただし、学習成果の利用状況については、まだ途上段階にあるのが現状である。学習成果を重視したものや、資質能力に重点を置いた取組はみられるものの、実際の利用状況としては、おそらく全大学で取り組んでいる国というのは少ない。特に、達成度評価については非常に僅かな国でしか実施されていないのである。

ところで、欧州委員会は教育等直接に権限を有しない政策領域に「開放型政策協調手法 (open method of coordination: OMC)」といった政策を採っている。各種指標を用いて各国の取組状況を比較しつつ政策実施を促すような手法であり、高等教育政策にも活用されている。他方欧州委員会は、ボローニャ・プロセスの準メンバーとして、諸活動に対する財政的な支援を通じて強い影響力を行使している。教育に関する指標として OMC に用いられているものとして、在籍維持率、理数系の大卒者数、生涯学習への参加数、人的資源開発投資などが挙げられる。OMC の採用は、高等教育における適格認定や大学ランキングの導入と軌を一にするものであって、各国を比較することによってリスボン戦略に合うような政策を実施するよう各国に促すのである。

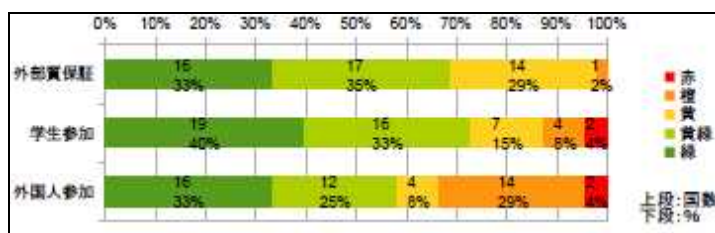
そうしたことから、ボローニャ・プロセスとリスボン戦略にある程度の収れんが見られる。ボローニャ・プロセスにおける質保証の枠組みがリスボン戦略の目的である世界最先端の知識経済をつくるといった目的にも活用されて、それと同時に高等教育の市場化の促進が図られるのである。そうした政策のあり方に対して教員や学生などからの反発も少なからず存在する。

以上の欧州の評価制度をまとめると次のようになる。一方で、ESG、EQAR、FQ-EHEA、さらには欧州単位互換制度（European credit transfer system: ECTS）など、全欧的な枠組みが整備されていることである。また、超国家的な組織である欧州委員会あるいは EUA（欧州大学協会）や ENQA といった団体・ネットワーク組織の活動があることである。このような収斂に向けた全欧的取組がある一方で、各国の制度に大きな多様性があること、伝統的な公的な性格もまだ維持されていること、学生参加を特徴とするような多様なステークホルダーが関与していることも重要である。そして、最近の政策の傾向の一つは、学習成果への着目、内部質保証の重視、適格認定（accreditation）よりも手順審査（audit）の尊重である。次にそれらについて言及したい。

## 第1節 質保証（評価）制度の展開

### （1）全欧的枠組みの適用

ここでは、いくつかの枠組みの適用状況を紹介したい。まず、欧州の約半数の国が自国の質保証制度を ESG に基づいて見直したという報告がある。ただし、項目によってはその適用に差が生じている。



出典：BFUG Working Group on Stocktaking（2009）

図1 ESGの項目別適用状況

図1は、左にいくほどESGに対応していて、右にいくほどあまり対応していないことを意味している。たとえば外国人参加はあまり図られていない国が少なからずあるし、また、学生参加も同様に図られていない国が若干残っていることが分かる。

次に、国内資格認定枠組みについてである。これは直接には評価制度とは関係しないが、評価を行う際、特にプログラムの評価を行う際の基準になるために無視できないものである。ただし、全面的に欧州の資格認定枠組みを国内に取り入れているのは僅かである。現在まで、蘭語圏ベルギー、ドイツ、アイルランド、オランダ、イギリス、スコットランド

の6カ国・地域が採用している。ボローニャ・プロセスの運営委員会（Bologna Follow-up Group: BFUG）による達成度調査（stocktaking）の評価基準は各回毎に厳しくなっており、実際2007年から2009年にかけて資格認定枠組採用の程度が下がった国が増えている。同時に、EUが生涯学習の枠組み（コペンハーゲン・プロセス）をつくっており、そのなかに高等教育の枠組みを全部取り入れようという試みがある。この試みは高等教育側への協議が十分になかったため、各国内で両枠組対立関係や混乱が生じてしまっている。EUAは、高等教育の枠組みが先行してできているなかで、EUの生涯学習の枠組みが高等教育の枠組みをあまり考慮していなかったことに批判的な意見を述べている。

なお、この質保証議論と密接に関わるのが学生移動の問題であるが、それ影響を与える前述教育構造同調事業（Tuning、以下「同調事業」）についても補足しておきたい。同調事業に関する取組は、まだそれほど活発ではない。実際同調事業が各国の大学で受け入れられているか、特に国内資格認定枠組みに反映されているかという点では、余り進んでいない状況である。流動性という観点からは、例えばECTSについては特にイギリスの大学はほとんど大陸の大学で得たECTSを認めていないし、まして外国の学修に基づいて学位を出すことを認めていないと聞く。したがって、確かに同調事業のような取組はあるものの、大学の現場ではまだまだ受け入れられていないし、国家全体としてもさほど資格認定枠組が整備されていない。今後これらの取組がどうなるかも、まだよくわからないところである。

とりわけ学習成果の取り扱いがまだまだ研究段階であり、とてもそれを測定するといった状況ではない。そのあたりの研究が進まないと、学習成果を明確にしてそれを流動性拡大に利用するところまでは行かないと思われる。

最後に、ボローニャ・プロセスで決められている学位附属書（diploma supplement）に触れたい。すなわち、学位を出す際にそれに付随する文書を出して、そこに学位を授与された者が学んだことを書くというものである。

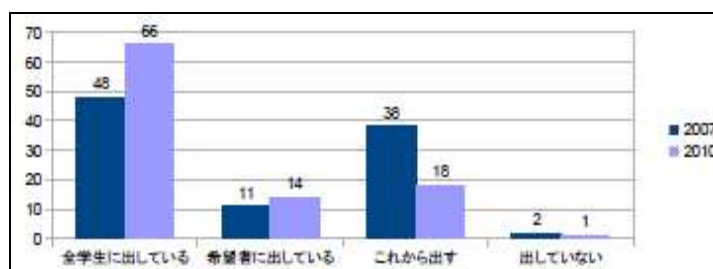


図2 学位附属書

図2のとおり、確かに発行する大学は増加しているものの、まだ発行していない大学もある。また、学位附属書に用いる言語は通用性がある言語という定めがあるものの、明らかに通用性がない言語で、つまり（小国における）自国の言語でしか出していない国も少なからずある。さらに、学習成果をそこに記載することになっているものの、学習成果につ

いては記さずにプログラムの内容を書くだけのものもある。そうした学位附属書も少なくなく、今のところ多くの国であまり活用されていないと聞く。

## （２）制度の多様性

はじめに、外部質保証のあり方をみてみよう。質保証機関の有無、質保証機関のあり方を示したのが図３である。



出典：Eurydice（2010）

図３ 国別の外部質保証の在り方（2009/2010）

質保証機関があつて助言的であるところ、質保証機関があつて監督的（認証権を有する）ところ、質保証機関を設置したが未運用のところ、質保証機関がないというところに分類される。質保証機関がないところは僅かだが、まだ若干残っている。また、フランスは助言的とされているが、確かに最終的な権限は既述の AERES ではなく政府が持っていて、法令上は助言的とは言えなくはない。しかし、実際は評価機関の評価は政府の意向に沿って実施されており、その結果が政府によってそのまま採用されている。そのため、AERES は制度上は助言的とはされつつも、実質的には極めて監督的機関に近いといえるだろう。

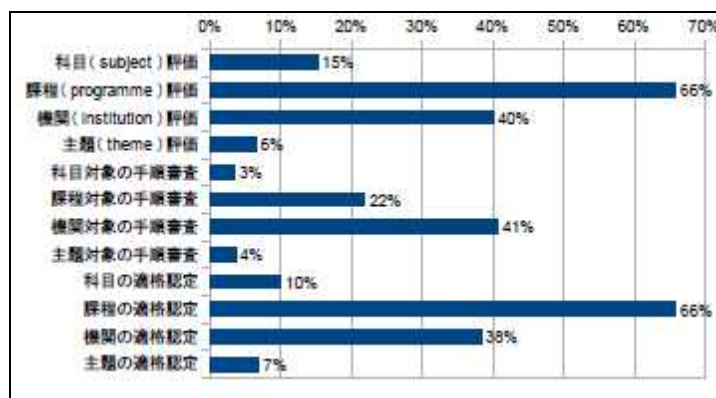
次に、図４は、科目（subject）、課程（programme）、機関（institution）の認証権限がどこにあるかを示したものである。機関、中央政府、地方政府、あるいは、質保証機関などとさまざまな対応がみられるところである。なお、ここで科目（subject）というのは、基本的には学問領域や特定の活動領域（職業教育の場合など）についての一連の授業等を意味する。課程（programme）は、米国でいうカリキュラムに相当し（欧州の言語では英語に訳す際に” curriculum” はあまり使わない）、一つの学位を出すことにつながるある程度体系化された教育活動である。たとえば、物理だったり、あるいは法学だったり、それを履修すると学位が出るといった一つの集まりがプログラムである。

|      | 科目<br>(subject) | 課程<br>(programme) | 機関<br>(institution) | 回答数 |
|------|-----------------|-------------------|---------------------|-----|
| 機関   | 25              | 21                | 5                   | 30  |
| 中央政府 | 9               | 16                | 21                  | 26  |
| 地方政府 | 5               | 11                | 16                  | 16  |
| QA機関 | 6               | 20                | 13                  | 22  |
| その他  | 3               | 6                 | 7                   | 7   |
| 別途記述 |                 |                   |                     | 20  |
| 回答数  |                 |                   |                     | 46  |
| 無回答  |                 |                   |                     | 5   |

出典：ENQA（2008）

図4 科目、課程、機関の認証権限者

そして、質保証機関による手法・対象別外部質保証活動の実施状況を示したものが図5である。



出典：Costes, et al.（2008）

図5 質保証機関による手法・対象別外部質保証活動の実施状況

科目、課程、機関、主題について評価している。評価というのはエバリュエーション (evaluation) を意味する。さらに、手順審査 (audit)、適格認定 (accreditation) についても、それらがどれだけ実施されているかを示している。たとえば、課程をみると、エバリュエーションを行うのが約3分の2である。同じ程度に適格認定が実施されている。そして、4割程度が機関の評価と同時に適格認定も行っている。おそらくこれは同じ機関が担当していると思われる。機関を対象とする、手順審査がいちばん多くなっている。

最後に、図6は外部質保証に関する各種要素の重要度を比較したものである。



| INDICATE THE RELATIVE IMPORTANCE TO YOUR AGENCY'S MISSION OF THE FOLLOWING WHERE THEY ARE OBJECTIVES OF THE MAIN TYPE OF EXTERNAL QUALITY PROCEDURES UNDERTAKEN BY YOUR AGENCY. | UNIMPORTANT | OF LITTLE IMPORTANCE | MODERATELY IMPORTANT | IMPORTANT | VERY IMPORTANT | NOT AN OBJECTIVE | RATING AVERAGE | IMPACT AVERAGE | RESPONSE COUNT |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|----------------------|----------------------|-----------|----------------|------------------|----------------|----------------|----------------|
| Transparency                                                                                                                                                                    | 0           | 0                    | 0                    | 7         | 37             | 0                | 4.84           | 4.84           | 44             |
| Quality improvement/enhancement                                                                                                                                                 | 0           | 0                    | 0                    | 11        | 33             | 0                | 4.75           | 4.75           | 44             |
| Accountability                                                                                                                                                                  | 0           | 0                    | 3                    | 14        | 24             | 3                | 4.51           | 4.20           | 44             |
| Accreditation                                                                                                                                                                   | 0           | 1                    | 2                    | 4         | 32             | 5                | 4.72           | 4.18           | 44             |
| Student (or prospective student) protection                                                                                                                                     | 0           | 1                    | 3                    | 10        | 22             | 8                | 4.47           | 3.66           | 44             |
| Estimation of graduate learning outcomes                                                                                                                                        | 0           | 2                    | 4                    | 15        | 17             | 6                | 4.24           | 3.66           | 44             |
| Make international comparisons with similar institutions/programmes                                                                                                             | 1           | 2                    | 8                    | 16        | 8              | 8                | 3.80           | 3.09           | 43             |
| Make national comparisons with similar institutions/programmes                                                                                                                  | 3           | 1                    | 11                   | 13        | 7              | 9                | 3.57           | 2.84           | 44             |
| Ranking                                                                                                                                                                         | 5           | 7                    | 4                    | 0         | 1              | 24               | 2.12           | 0.88           | 41             |
| Other                                                                                                                                                                           |             |                      |                      |           |                |                  |                |                | 7              |
| answered question                                                                                                                                                               |             |                      |                      |           |                |                  |                |                | 44             |
| skipped question                                                                                                                                                                |             |                      |                      |           |                |                  |                |                | 7              |

出典：Costes, et al. (2008)

図6 外部質保証に関する各種要素の重要度

いちばん左の列に、上から外部質保証の一つの観点である **Transparency**、**Quality improvement/enhancement**、**Accountability** などがある。それらがどれだけ重視されているのかをこの表は示している。真ん中ぐらいいある **VERY IMPORTANT** というところをみると、たとえば **Transparency** が 37、その次の **Quality improvement/enhancement** が 33、一つ飛んで **Accreditation** が 32、このあたりが多く、**Accountability** は 24 と半数程度であって、必ずしも **Accountability** というのは重視されていないという印象を与える。

### (3) 公的性格と学生等の参加

欧州の多くの国では、依然として大学は国立が主流である。フランスはもちろんのこと、デンマークやアイルランドもまた大学はすべて国立であるため、質保証についての政府の地位や役割は非常に重要である。ほとんどの質保証機関は法令で位置づけられており、約 4 割の機関の理事会や評議会に政府代表が席を占めている。およそ 8 割程度の質保証機関は政府の資金に依存していて、そのうち半分以上は全面的に政府による資金で賄われている。国への依存率が非常に高いのである。ただし、反対に、まったく国へ依存しない機関が 15%ほどあって、おそらくこれらはビジネスのような比較的收入が潤沢な領域別評価機関であろう。

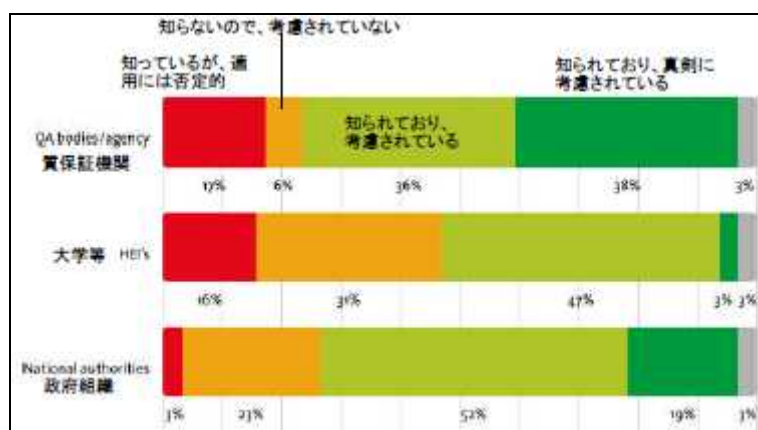
そして、伝統的な制度である視学制度が残っていることも特徴である。フランスには、既述の評価機関と合わせて、国民教育研究行政監査総局といわれる大臣直属の組織がある。98 人の視学官を抱えていて、教育研究——初等・中等教育も含む——を対象とした視学を実施して、教育機関に直接助言をしたり、担当大臣への報告書を提出したりする活動を行っている。大学の評価も含んでいて、組織のすべての機能が対象となっている。そのため、単純に手順審査を実施するだけではなく、実際にその活動から得られる生産物を評価した



り、個別ではないものの職員を評価したりするのである。

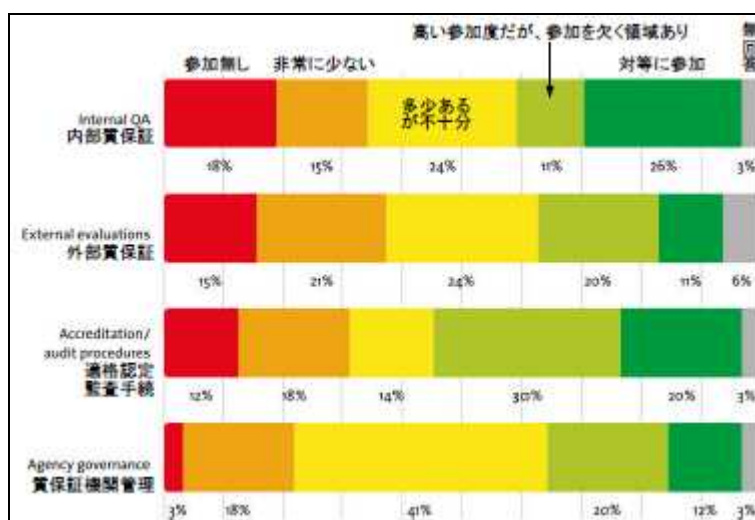
この視学制度は、オランダでも残されている。教育視学局——*Inspectie van het Onderwijs*、英語にすると *Inspection of the Education*——があつて、かつてはフランスと同じように機関の評価活動を監督していた。そのため、オランダではメタ評価を視学制度のなかで実施していたのだが、前述の NVAO ができた段階で教育視学局の審査対象が NVAO の評価活動に移り、その適格認定全般を監視——*toezicht* と表現されている、英語では *supervision*——することが教育視学局の職務になっている。

図 7、図 8 は、ESU（ヨーロッパ学生連盟）という全欧の学生団体の連合会が調査した、質保証活動にどれだけ学生が参加しているかを示すものである。



出典：ESU（2009）

図 7 学生参加にかかる ESG の適用状況（機関別）



出典：ESU（2009）

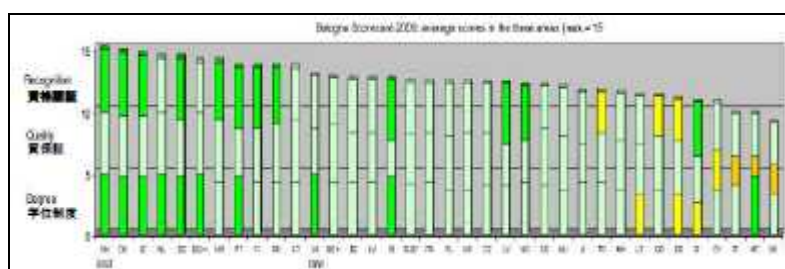
図 8 学生参加にかかる ESG の適用状況（活動領域別）

左側になるほどあまり参加していない、すなわち ESU によって評価されていないことを

示し、右側へ行くほど学生が参加している、評価が高いということになる。国別では、どちらかというと北のほうで概して参加率が高く、南のほうは、例外的にスペインの参加が高いことを除けば概ね低い。

#### （４）超国家組織の活動

欧州委員会が開放型政策協調（OMC）を推進していることは先に述べたとおりである。この枠組の下で、委員会は既存の様々な指標を収集して、各国の業績を比較している。図9は、質保証にかかるボローニャ・プロセスの推進状況を独自の基準で点数化し、合計した点数が高い順に並べた結果である。



出典：CEC（2011）

図9 ボローニャ・プロセスの業績に基づく評価点の順位

後述のようにどの国がトップということはあまり意味がないのだが、結果だけを見れば、左側の点数の高い位置にスコットランド、デンマーク、オランダ、スウェーデン、蘭語圏ベルギーがあり、右側へ行くほど成績が悪い国であり、いちばん右側が最も点数の低いスロバキアである。しかし、欧州委員会が利用している指標はボローニャ・プロセスの進捗状況報告書（stocktaking report）にある評価結果を点数化したものであることの留意する必要がある。すなわち進捗状況報告書自体は5段階評価をしているだけで、それぞれの評価の点数化をまったくしていない。また、この図のように全ての事項を総合化した評価というものも行っていない。つまり、進捗状況の評価には多様な側面があって、それぞれの項目ごとにどれを重視するかについては各国の判断に任せている。しかし、欧州委員会は5段階評価の各項目を点数化して全部合わせるといふ、非常に単純で乱暴ともいえるような集計を行った上で国の順位づけを行っているのである。

全体的に OMC でとられている手法にはこうしたものが多い。非常に単純化されているのが特徴の一つであって、精緻化された比較というには程遠いといって差し支えないだろう。取り上げられている指標のうち質保証にかかるものは限定的であって、まして学習成果についてはまったくといっていいほど取り上げられていない。一部例外的なのは外国語の操作能力や情報操作能力であるが、しかし、これらにしても第二外国語をどれだけ勉強しているかとか、それらの程度は高いか低いとか、その程度について聞いているだけであって、細かいところまで見ていないのが現状である。

続いて、EUA の機関評価を確認しておこう。これ自体は手順審査型の評価である。その目的は、機関の戦略的経営と内部質保証文化の継続的発達を促すことにある。自己評価に基づいて訪問調査を行い、期間とのやり取りを踏まえて評価報告書をまとめ、最後に勧告を出すといった手法がとられている。すでに 20 年近い伝統があつて、これまで 45 カ国 260 機関——日本の機関も含まれている——を評価したということである。他方、手順審査実践上の課題も指摘されている。たとえば、透明性やコミュニケーションを伴わない執行部は課題として認識されている。また、構成員の積極的な参加や質保証の仕組みの分権化は必要事項である。好業績への報償が必要であると同時に、低業績への制裁の回避、つまり、総括的で制裁的な評価は避けるべきであるということがいわれている。さらに、学問領域ごとに手法の違いというものも尊重すべきであるといったさまざまなことが、これまでの経験から明らかになっているのである。

## 第 2 節 教育研究の質の向上に向けて

### (1) 質保証の重点——内部質保証

質保証の重点は、近年、内部質保証に置かれるようになっている。EUA などには特に内部質保証の発達に重点を置いて活動しており、さまざまな研究会を開催し、多様な調査研究活動を行っている。その結果、内部質保証は機関の自律性と密接に関係することが明らかになった。すなわち、内部質保証の発達を促すためには機関の自律性が不可欠であつて、自律性が低いと内部質保証も発達しない。さらに、大学内部においても次第に中央集権的な運営のあり方から部局への分権的なあり方へ移行していて、質保証自体も分権化すべきと考えられている。他方、さまざまな課題も残されていて、たとえば、学習支援が内部質保証と密接に関係する、その発達が必要であるといわれているものの、ここ近年の間、学習支援への取組状況はほとんど変わっていないのである。いずれにせよ大学では、内部質保証活動の強化が、ここ近年の間での変化のうち最も大きな変化の一つであると捉えられている。



図 10 教育課程評価の手法の実施状況 (%、TRENDS2010)

図 10 は、教育課程評価の中身を示したものである。いちばん高い率で挙がっているのが、教育課程の内容、教授法、達成すべき学習成果が外部の適格認定等によって評価されている項目である。ただし、このことは適格認定等で評価されることが依然として主流であることを意味する。また、2007 年から 2010 年にかけて教育課程や研究活動等の項目が若干

減っているが、それは、学内において分権化されていることもその一環ではないかと考えられている。

次に、EUA の「高等教育変革課題のための質保証」研究事業の勧告を確認したい。質保証は文脈に依存的であることに配慮すること、大学の変化能力の発達を促すような開発的手法を採用すること、質保証の諸手順は関係者全体——特に、学生——を巻き込む包括的なものであること、主要関係者の関与・対応能力を適切に確保すること（職員開発や学生への支援・訓練を含む）、大学と質保証機関間の連携の強化・拡大、危険を冒すことや失敗に対する寛容性——評価は形成的であるべきだ——、そして、優良実践の共有が勧告されている。

## （２）質保証を超えて

「質文化」の普及は重要な課題とされている。適格認定はいまだに主流であるものの、大学においては、なおそれに対する拒否感が強い。たとえば、蘭語圏ベルギーでは、従来からピア・レビューをもとにした外部評価が存在していたが、それに加えて、2000 年頃に適格認定が導入された後に、広範囲にわたって評価項目が精緻化された。その結果、大学は自己評価報告書をつくるのに膨大な作業に追われるといった状況になり、その結果目先のことで手一杯で、戦略を考えたり改善を図ったりすることが困難になった。同時に、評価側にしても評価者が自己評価報告書を全部読むのが大変で、評価の実施が困難になってしまった。そこで、近年では簡素化された制度に変革された。ただし、だからといって、適格認定そのものが放棄されたわけではない。以前のピア・レビューに基づく外部評価は形成的な評価であったものの、報告を受けた後に引き出しのなかに放っておかれたり、実際はほとんど考慮されなかったりすることもあって——フランスの以前の大学評価委員会、今の AERES の前の CNE の時代のように——、そうした事態を懸念する政府は適格認定を今後とも維持したいという意図を持っている。当然、大学のほうでは、このような政策の評判は極めて悪い。

一部の国々、たとえば、スコットランド、フィンランド、あるいは、スイスといった国は手順審査へ移行している。渡邊（2009）によれば、フィンランドの手順審査（オーディット）の目的は、機関みずからがその活動について設定した質に関する目標を明確にすること、機関が教育等の活動の質を維持・向上させるために、どのような手法、どのような手順を用いているかを評価すること、機関の質保証システムが意図されたとおり機能し、みずからの運営改善のために有益な情報をもたらし、効果的な改善手段となり得ているかを評価することにあるという。ただし、こうした動向は一部の国のものでしかなく、また、適格認定そのものも非常に多義的な行為なので、手順審査と適格認定の間には、さまざまな段階で競合があるというのが現状である。

## おわりに

欧州における評価制度は多様である。政策の方向としては、学習成果を重視する方針が示されているものの、それを測定したり、実際に流動性拡大等に活用したりするといったことまでには至っていない。一部、北欧諸国はある程度進んでいるようではあるものの、それでもごく一部の領域（例えば資格につながるようなもの）を除いては、成果の測定までには辿り着いていないのである。また、外部質保証から内部質保証の重視への転換、さらには、学内における分権、質文化の定着と質向上への動きもみられる一方で、依然として適格認定を中心とした一定の基準に基づいた評価活動も根強く存在している。

リスボン戦略が目指しているように、大学進学率並びに大学修了率の向上を図る政策を各国は採っている。しかし、それと同時に学習成果の保証を追求するということは実際には困難であって、こうした取組が欧州で成功しているとはいえない。基準を定めて、これに基づいた適格認定等が行われるのだが、それは大学の自律性と抵触するということが現場で問題視されている。他方、リスボン戦略が依存する人的資本論が実証されていないという批判も強く、高等教育の学位保持者の増加が経済成長につながることも自体が立証されていないままである。

資質能力を育成できるかどうか、学習成果が測定できるかどうかについても疑問があるとされていて、資格認定の枠組みも十分に普及されているとはいえない。たとえば、前述の教育構造同調事業のような作業も進んではいるものの、現場までにはまだまだ普及していない。評価を行う側もまた評価能力が問われており、これをいかに育成するかが課題である。また、学生を参加させるという方向性が示されているにもかかわらず、なかなか学生をうまく活用できていないといった問題も残されているのである。

最後に、欧州と比較して、日本の質保証に関する議論の「狭さ」について、一言述べておきたい。日本の質保証についての議論は、欧州のそれと比べると、評価のための技術に傾斜している感じが否めない。マネジメントの手法（PDCA など）をいかに導入するかという観点や、管理者がどうやって質保証をしていくか、またそのための FD・SD をすすめるという発想が非常に強いように思われる。それに対して欧州のほうは——全部が全部うまく行っているわけではないが——、利害関係者をいかに参加させるかが重視されており、特に学生や教員を取り込んだ仕組みを作ろうとしている。いろいろ難しい問題はあるにせよ、評価制度が形骸化しないための一つのポイントはそこにあるのではないか。組織論の観点からも、利害関係者を入れて学生などが参加するような組織をつくっていけば、ある程度質保証の活動が実効的に機能する可能性が高まるのではないか。欧州でいわれる「質文化」という言葉にもそのような考えが反映されているように思う。日本でも、技術以外の点がもっと重視される必要があろう。

## 主要参考文献

- BFUG Working Group on Stocktaking, 2009, *Bologna Process Stocktaking Report 2009*, Leuven/Louvain-la-Neuve: Report to the Bologna Process Ministerial Conference 2009.
- Commission of the European Communities, 2011, *Progress towards the Lisbon objectives in education and training: Indicators and benchmarks 2010/2011*, Brussels: EU.
- Costes, N. et al., 2008, *Quality Procedures in the European Higher Education Area and Beyond • Second ENQA Survey*, Helsinki: ENQA.
- ESU, 2009, *Bologna with Student Eyes 2009*, Brussels: ESU.
- Eurydice, 2010, *Focus on Higher Education in Europe 2010: The Impact of the Bologna Process*, Brussels: EACEA.
- Hofmann, S., 2005, *10 Years on Lessons Learned from the Institutional Evaluation Programme*, Brussels: EUA.
- Kalkwijk, J., 1992, “The Netherlands: The Inspectorate Perspective,” A. Craft (Ed.), *Quality Assurance in Higher Education: Proceedings of an International Conference Hong Kong, 1991*, London: Falmer Press, 97-107.
- 大場淳、2009、「ボローニャ・プロセスと学生参加」日本教育行政学会研究推進委員会編『学校と大学のガバナンス改革』教育開発研究所、224-241。
- 大場淳、2011、「欧州における高等教育質保証の展開」広島大学高等教育研究開発センター編『大学教育質保証の国際比較（戦略的研究プロジェクトシリーズIV）』広島大学高等教育研究開発センター、1-24。
- 大場淳、2011、「知識基盤社会と大学教育—欧州における取組から—」広島大学高等教育研究開発センター編『知識基盤社会と大学・大学院改革（戦略的研究プロジェクトシリーズ）』広島大学高等教育研究開発センター、予定。
- 渡邊あや、2009、「フィンランドの質保証システムとボローニャ・プロセスへの対応」広島大学高等教育研究開発センター『欧州の質保証とその政策過程に関する研究会』配布資料。