

【研究論文：依頼原稿】

大学評価・学位授与機構による試行的大学評価事業の評価

齋藤 貴浩 林 隆之
 大学評価・学位授与機構 評価研究部
 tksaito@niad.ac.jp

要 約

大学評価・学位授与機構（以下「機構」と記す。）は、大学評価を行う第三者評価機関であり、2000年度から2003年度まで国公立大学及び大学共同利用機関を対象に、延べ550組織に及ぶ試行的評価を実施した。評価とは、政策、施策、事業等を体系的かつ客観的に分析し、価値判断を行うことであるが、しかし、その評価自体も、各種資源を投入して評価対象並びに社会に対して何らかの変化を生じさせる事業である。つまり、評価自体も一つの事業として評価の対象となりうる。本研究では、機構が行った試行的大学評価について、ロジックダイアグラムを基に評価事業の設計と評価結果の妥当性並びにその成果について論じ、また費用効果の観点から評価事業の効率性を検討する。さらに、新たな大学評価システムの構築に活かすため、試行的評価経験と、評価の評価によって新たに得られた教訓から提言を行う。

キーワード

大学評価、メタ評価、ロジックダイアグラム、費用効果、自己評価

1. はじめに

ここ数年というごく短い期間で、初等、中等、高等教育の別を問わず、教育への第三者評価の導入が急速に進んでいる。教育は誰もが体験し、また身近な存在であるがために、教育の評価は容易であると思われがちである。しかし、実際にはこれほど評価が困難な事業はない。例えば、教育の成果は多様であり、目に見えず、個人に依存し、広く波及し、長期的な展望を必要とし、そして極めて広範な利害関係者が相反する利害を有しつつ強く主張もしている。さらに、大学を評価すると

なれば、学術・研究という高度な専門性とその特殊性を考慮に入れる必要がある。このような特殊な事業評価を、どうすれば誰もが納得するような形で遂行できるのかという命題は、最適解があるのかどうかすらわからない。

教育に限らず、これまでに評価関係者は評価を通じてどのように政策や事業をより良くしていくかということに努力してきた。しかし、評価という事業自体もまた、人的資源や資金を必要とし、事業や組織、さらには社会を良くしていくための一つの事業に他ならない。事業である以上は、改善や説明責任の遂行を目的とする評価という活

動、それ自体にも、改善の必要や説明責任が生じている。実際に、大学の第三者評価を目的に創設された大学評価・学位授与機構（以下「機構」と記す。）による試行的大学評価では、評価が大学に大きな負担と影響とを与えることから、評価対象である大学側から常に「メタ評価（評価の評価）」の実施を求められてきた（例えば国立大学協会2002, 2003a, 2003b）。

「評価の評価」を意味するメタ評価は決して新しい概念ではない。例えば、大川（2002）はODAプロジェクトの評価をメタ評価し、妥当性、論理性、費用効果の視点の欠如などを指摘している。また国際協力機構外部有識者事業評価委員会/2次評価作業部会（2006）は、3回目の2次評価結果として公表されており、多くの事業評価報告書を客観的かつ定量的に評価し、共通する傾向や特徴的な傾向を導き出すことによって国際協力機構の評価の質の改善に寄与している。しかし、これらの研究は主として「評価が適切に行われたか」という評価にとどまる。つまり、事業評価として見た場合には、活動、あるいはプロセスの評価であって、アウトプット、アウトカムの評価ではない。評価事業を他の事業と同様に評価するためには、評価の目的である「改善を通じた質の向上」や「説明責任を果たす」という、アウトカムを評価しなければならないのは当然である。

本研究は、機構が過去に行った試行的大学評価を1つの事業としてとらえ、そのアウトプットやアウトカムの評価を行うことにより、優れた点と問題点とを明らかにするとともに、試行的評価で得られた経験と教訓とを新たな評価システムに活かすことを目的とする。

2. 大学評価・学位授与機構による試行的大学評価の概要

(1) 試行的大学評価の背景

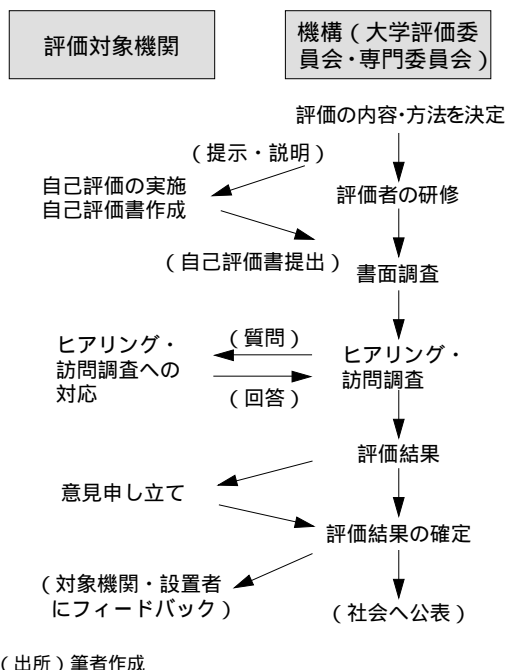
大学の評価は、1991年の大学設置基準の大綱化に伴い、自己点検・評価が努力義務化されたことで大きく展開した。つまり、大学に対する法的規制を弾力化して大学の自主的かつ自律的な改革を

図る一方、評価を通じて自らの努力で質の向上を行うよう促したのである。多くの大学が自己点検・評価を実施したものの、実際には形式的評価に陥り、改善に十分結び付かない例も多くあった。そのため、大学審議会（1998）は、自己点検・評価の充実を図るとともに、第三者評価システムの導入などを通じて多面的な評価を行い、大学の個性を伸ばし、教育研究の内容・方法の改善につなげるシステムを確立する必要性を示した。これに基づき、1999年には自己点検・評価が努力義務から義務とされ、さらに外部評価が努力義務とされた。そして2000年には、大学の第三者評価を行う機関として大学評価・学位授与機構が改組によって設立された。

機構は、本格的な第三者評価の実施を2004年度に設定し、当面は大学評価に関する経験を蓄積するため、学問分野・機関数を絞って段階的に大学評価を実施することとした。この試行的評価は、2000年度から2003年度まで、設置者から要請のあった国立大学及び大学共同利用機関（以下「大学」と記す。）を対象とし、全学テーマ別評価、分野別教育評価、分野別研究評価の3つのテーマに分けて、3回のサイクルで実施された。全学テーマ別評価では評価対象は大学全体であり、主としてすべての対象機関の協力のもと、「教育サービス面における社会貢献」、「教養教育」、「研究活動面における社会との連携及び協力」、「国際的な連携及び交流活動」の4つの評価が実施された。分野別教育評価、同研究評価では学部または該当組織を対象に、理学、医学、法学、教育学、工学、人文学、経済学、農学、総合科学の9分野を対象とする評価が実施された。4年間の実施期間で評価を実施した機関は122機関、延べ550組織に上った。

2004年度から予定されていた第三者評価の本格実施は、学校教育法の改正及び国立大学法人法の制定等により実際には行われなかった。しかし、機構は、大学、短期大学、法科大学院、高等専門学校を対象とした認証評価機関の一つとして2005年度から認証評価を実施しており、また国立大学法人評価に際しても国立大学法人評価委員会の要請を受けて教育・研究活動に関する評価を実施する予定である。試行的大学評価で得られた教訓は、これらの評価の基礎として活かされる。

図1 大学評価のプロセス



(2) 試行的大学評価の方法

機構による試行的大学評価は、「大学の教育研究水準の向上に資するため、大学の教育研究活動等の状況について評価を行い、その結果について、当該大学及びその設置者に提供し、並びに公表すること」を目的とし、実施された。

図1は機構の行った試行的評価のプロセスを示したものである。評価は評価対象機関による自己評価と、訪問調査あるいはヒアリングが基礎となる。最初に、機構において大学評価委員会、各専門委員会が評価の内容・方法を決定し、評価対象機関に提示、説明を行う。評価対象機関はそれに基づいて自己評価を行う。専門委員会は提出された自己評価書に関して書面調査を行い、不明な点等についてヒアリングまたは訪問調査で確認し、そして評価結果が作成される。なお、評価に際しては、機関別に編成された評価チームや、学問分野別に編成された部会が実際の評価にあたる。評価結果は公表の前に各評価対象機関に通知され、意見申し立てを経て評価結果が確定される。最終

的に、評価報告書は対象機関・設置者にフィードバックされることにより、対象機関における質的向上を促進することが期待され、また社会に対する公表によって、アカウンタビリティの促進が期待されている。

3. 分析方法

本研究で分析対象としたデータは、試行的評価の中であらゆる機会で作成してきた(1)意見調査の結果に加え、今回の検証作業で改めて(2)質問紙調査、(3)インタビュー調査を実施し、収集したものである。

(1)意見調査：機構が評価実施の際に、各実施段階において評価対象機関、評価者、関係団体に行った意見照会の結果である。調査方法は、主として自由記述形式の質問紙を用いての郵送調査であり、文章単位で5,000以上の意見が収集された。それらの意見を取りまとめた結果、およそ500の意見に集約された。なお、詳細な結果は紙面の都合で省略しているが、研究評価については林ら(2003)、全学テーマ別評価については齋藤(2004, 2005)、ならびに試行的評価に関する検証委員会(2004)を参照されたい。

(2)質問紙調査：質問紙は(1)の集約された意見を参考に作成され、調査は平成16年7月にすべての評価対象機関に対して郵送法で行われた。この調査では、(1)の調査でカバーされていない、大学内の評価の過程や評価の成果等についての情報収集を行った。質問紙への回答は、各機関内での自己評価を実際に担当した方にお願ひし、回収数は539(回収率98%)であった。

(3)インタビュー調査：地域や規模等を考慮し、9大学及び1大学共同利用機関に協力をお願いし、平成16年9月から10月にかけてインタビュー調査を実施した。調査協力者は、評価担当者及び大学の運営に携わる方々であり、(1)(2)で十分に把握できなかった大学内の事柄について調査を行った。

(2)(3)の調査及び検証作業は、機構の試行的評価の当初の実施計画に基づき設計されたロジックモデルに基づいて行われた。ロジックモデルとは、プログラムの要素(資源)活動、アウトプット、アウトカムについて、またそれらの因果関係について明確なモデルとして示し、プログラムの問題解決に役立てる手法である(Bickman 1987, Wholey 1983, 1987)。実用に際しては、左から右へ、因果関係を同定した図を用いることが多い(例えばW.K. Kellogg Foundation 1998, 2004)。ロジックモデルは、問題解決にどの点を重視するかによって、セオリー・モデル、アウトカム・モデル、アクティビティ・モデル(あるいはそれらの複合モデル)と分類されたり(W.K. Kellogg Foundation 2004)。モデルの使用方法に関しても、プログラムのライフスパンの中で、フィージビリティ評価のため、改善のため、モニタリングシステムの開発のため、そして教訓の蓄積のためと、様々な目的で使用されたりする(Savaya and Waysman 2005)。また、ロジックモデルという名称や形式以外にも、プログラム・セオリーを使用するという点を中心に、多くのバリエーションが存在している(McLaughlin and Jordan 1999)。どの方法においても、プログラムを因果関係で捉え、問題点を抽出し、プログラムをより良いものにするという点は変わらない。そして、ロジックモデルの形式は変更不可能な決まり事ではなく、評価対象となる事業を明確に表す上での適切なモデルの選択は評価者に委ねられる。

本研究では、科学技術政策評価で定評のある英国マンチェスター大学のPREST(Policy Research in Engineering, Science and Technology: 工学科学技術政策研究所)の提唱するロジックダイアグラムを基礎として、プログラムのモデル構築を行った。このロジックダイアグラムでは、上下の階層構造で上位レベルの目的、下位レベルの目的、活動、期待されるアウトプット、アウトカム、インパクトの順に要素を並べ、それらの因果関係をリンクさせる。目的と期待されるアウトカムを別の階層とすることにより、目的には明示されにくい効果、負の効果、スピルオーバー的な効果についても明確に示されることを特徴とする(林 2002)。本研究でロジックダイアグラムを用いる目的は、試行

的・大学の評価を一つの事業として捉え、今後の大学評価のための論点整理とともに教訓を導くことにある。その点、PRESTのモデルは主たるプログラムの目的のみならず、さまざまな成果を考慮することができる点でふさわしいと考えられる。なお、ロジックダイアグラムの構築に関しては、試行的評価が当初の評価設計通りに遂行されたか、またその遂行に際して当初の想定と異なる点がないかどうか等を検討することが重要であるとの考えのもと、当初の評価の設計(基本的枠組み)をそのままモデルに組み込んで分析を行った。

以下に示すような45の要因を設定し、因果関係に基づき構築されたロジックダイアグラムは後の図6に示すとおりである。

- ▶ 目的(2): 大学の質の向上とアカウンタビリティ
- ▶ 目標(5): 目的を達成するための具体的な課題(必須とする要素)
- ▶ 基本的枠組み(9): 評価手法として容れなければならない要因
- ▶ 投入(評価の体制)(機構3、対象機関1): 評価、改善、公表の各体制
- ▶ 活動(評価の実施)(機構4、対象機関1): 評価に伴う各種の活動(評価方法を含む)
- ▶ 結果(5): 評価結果、評価経験の蓄積と改善、作業負担、マスメディアによる報道
- ▶ 成果(対象機関8、社会5): 評価経験の蓄積、活動の改善、社会からの理解等の各種成果
- ▶ 目的の達成(2): 大学の質の向上とアカウンタビリティ

4. 分析結果

分析は、試行的評価のインプット、プロセス、アウトプット、アウトカムのすべてについて、網羅的に実施した。その結果、「目的」、「目標」の構成、一連の「評価の基本的な枠組み」に関してはおおむね良好な結果であり、一部の要素を除いて設計通り実施されたという結論を得た。評価を行う機構側の「活動」の問題点として、評価の目的や方法の理解に関する問題等が挙げられたが、試行的評価の進行とともに改善された。一方、評

価の対象機関側の「投入」と「活動」(図6:中央右側) 一連の「結果」及び「成果」(図6:下段)は、その一部または全体的に、問題が生じていたことが明らかとなった。

そのため、ここでは対象機関側の「投入」と「活動」、つまり自己評価のプロセスと、試行的評価全体としての「結果」と「成果」の分析結果について論ずる。

(1) 自己評価のプロセス(対象機関側の「投入」と「活動」)

質問紙調査の結果、多くの機関では自己評価を担当する組織として委員会形式の組織が挙げられた。これら組織は約9割が常設の組織であり、既存の自己点検・評価の体制で試行的評価に対応していたことが明らかとなった。そして、実際の作業のために、評価担当組織の下にワーキンググループや部会などを設けている機関が多くを占めた。インタビュー調査では、当初は委員会組織で対応していたものの、実際には対象テーマに関する専門性が必要となることからワーキンググループ等を設置した機関もあり、過年度の評価経験を踏まえて適切な自己評価体制を模索した過程がうかがえる。また、中には評価に特化した常設の専門組織を設置している機関も見られた。インタビューでは、評価専門部署を設けることで、より全学的な組織的連携の下で評価を実行、統括できるため、問題点の把握等も認識しやすくなったとのことであり、この試行的評価の経験が昨今の国立大学内での評価関係組織の充実に繋がっている。

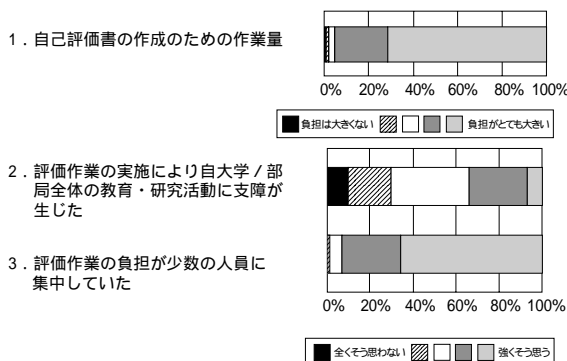
自己評価等の作業量については、意見調査では評価担当者に過大な負担が強いられているという意見がきわめて多くあった。図2は自己評価担当者の作業負担に関する質問紙調査の結果であるが、「1.自己評価書作成のための作業量」について、「負担が大きい(5段階評価のうち4と5)」とする回答が9割を超えており、評価負担の軽減は必要不可欠と言える。しかし、「3. 評価作業の負担が少数の人員に集中していた」という質問に対する肯定的意見も同じく9割を超えていることから、負担が大きい理由として、評価作業が日常の

業務に位置づけられていないこと、そして評価作業が一部の部署・人員に集中しているという評価対象組織内の要因が挙げられる。

(2) 評価結果の適切さと公表方法(「結果」)

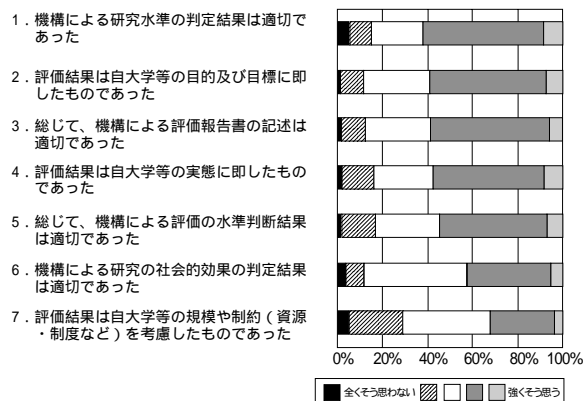
試行的評価の評価結果は、大学の目的に即して、評価項目ごとに判断結果を定型句で表現するとともに、優れた点と改善を要する点を評価報告書の中で指摘した。評価結果は試行的評価事業の評価の中で最も配慮すべき事柄の一つであるが、この評価結果(評価報告書の内容)の適切性について、評価対象機関の回答を表したものが図3である。

図2 自己評価担当者の作業負担



(出所) 筆者作成

図3 評価結果の適切性



(出所) 筆者作成

この結果から、評価結果は評価対象機関からおおむね肯定的に受け入れられている。判定結果、報告書の記述とも大学からの評価は悪くはない。しかし、「6.研究の社会的効果」、「7.自大学の規模や制約（資源・制度など）の考慮」については肯定的回答の割合が相対的に低く、比較的困難な社会的効果の慎重な扱い、そして各大学の特性をより考慮した評価が求められている。

これらの評価結果が「成果」を生むためには、それが関係者に伝わらなければならない。対象機関に対しては直接通知するために問題はないが、社会に対しては評価機関が結果を公開するだけでは周知に限界があるため、ロジックダイアグラム上、評価の「活動」と社会における「成果」とを結ぶ「マスメディアにおける報道」の役割が重要となる。

しかし、意見調査において、評価結果の公表方法に対する非常に多くの批判が寄せられた。特にマスメディア等による評価結果の取扱いについての批判である。評価結果は目的に即したものであるため、結果の比較は意味を持たないという認識を評価者側と評価対象機関で共有し、評価を実施し、記者発表の際にも十分に説明をしたにもかかわらず、評価結果がランキングの形で報道されたり、大学間の比較をされたり、あるいは良い結果を得た項目は報じられず悪い結果のみがクローズアップされるなど、大学評価の意図とその報道のされ方には大きな隔たりがあった。つまり、そのような評価結果の表し方は、現状の報道には適さないということであり、さらには社会の捉え方とも乖離しているということである。わかりやすい評価という意味では、その要求に答えられていなかった点を見過ごすことはできないし、逆に社会にランキングではない大学評価手法の説明をする必要もある。メディアが大学評価に果たす役割は、機構の評価が最終的に判断を社会に委ねる以上、機構と大学の利害関係者とを結ぶ意味で極めて重要であり、この関係は新たな評価制度でも変わらない。評価の趣旨・ねらいやその評価結果のとらえ方等、さらには大学の実際の状況の理解について、社会からよりの確な理解が得られるよう、社会に対する、機構、大学双方の一層の努力が求められる。

(3)大学の質の向上(対象機関で発現する「成果」)

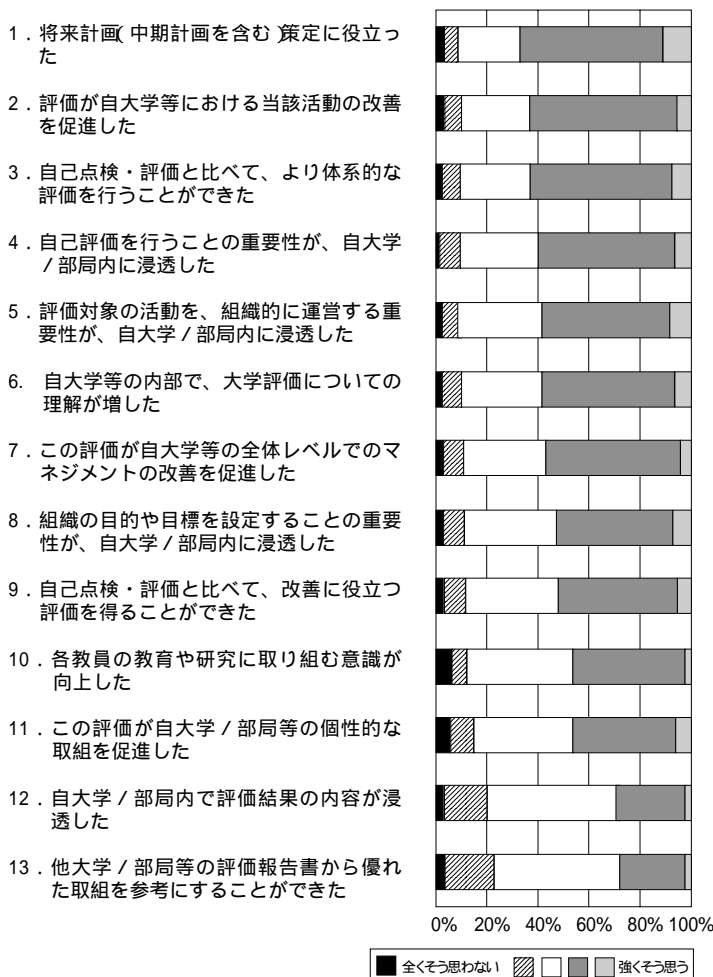
大学の質の向上を実現するために、試行的評価の基本的枠組みでは「大学の主体的な改善を促す評価」、「大学の個性を伸ばす評価」が目標とされた。これらの目標の達成について質問紙調査で確認したところ、その結果は図4のようになった。「2.当該活動の改善を促進した」、「5.組織的に運営する重要性が自大学内に浸透した」等、6割以上が肯定的な回答であり、自己点検・評価との比較を求めた項目も肯定的回答である。当初の設計通り、評価が改善を通じての質の向上に繋がっていることが見て取れる。

これは、自己評価よりも第三者評価が優れているということの意味するものではない。インタビュー調査では、自己評価によって自大学の課題を把握することができたというような意見が多くあり、評価の一環として、自組織に都合の良い自己点検・評価ではなく、外部の者によって決められた項目で実施した自己評価作業自体が、大学の改善に向けて有効に機能したものと考えられる。また、大学評価が、その後の中期計画の策定、マネジメントの改善や、評価の重要性の認識等にも繋がっていることは、その効果が多様であることを意味している。

さらに、評価後の具体的な改善事例についての記述を求め、それらの各改善事例について機構の評価がどの程度参考となったかを質問したところ、全対象機関122機関中の109機関(89%)から、のべ1,024件の回答が得られた。その結果をまとめたものが表1である。

回答のあった改善事例のうち、全評価区分を通じて多かった事例としては、学内組織の連携強化/整備に関するものや、自大学内での評価体制の整備や改善システムの強化(体制整備、根拠資料等の収集の経常化、アンケート等による問題点の把握など)、周知公表の体制(活動)の強化・充実など、教育研究活動等を実施する上での基盤的な部分の更なる改善に関する事項が挙げられる。これらについて、評価の参考度が4(ある程度参考となった)以上であるものは825件(81%)となった。このことから、改善の促進に関して試行的評価の一定程度の貢献が認められる。

図4 大学の質の向上



(出所) 筆者作成

しかし、これら改善の促進は、評価結果を受けての改善のみならず、試行的評価への組織としての対応、また自己評価の過程が直接的に改善を促したとする意見がインタビュー調査では多くあった。例えば、既に自己点検・評価や日常の活動の中で認識されていた課題に関して、試行的評価を契機に積年の課題への取り組みが一挙に進んだというコメントもあり、機構の評価が改善の促進装置として働いたことが指摘できる。これは当初の想定とは異なる因果関係となっている。さらに、「評価対象の活動を組織的に運営することの重要

性が自大学内に浸透した」などのマネジメント上の意識改革の効果についても半数以上の肯定的回答が得られており、大学の改善への間接的な効果も少なからず存在したと言える。

一方、個性の伸長については肯定的な回答が4割程度にとどまった。このことは、大学の多様性を尊重するために採られた「目的に即した評価」という評価の枠組みが、個性を犠牲にしないためには機能しても、個性の促進のためには十分ではないことを示している。また、試行的評価を実行可能なものとするために、項目、要素、観点等の

表1 改善事例の件数と評価の参考度

評価区分	改善事例	1	2	3	4	5	記述なし 参考度の 総計	総計	(回答者数)
		全く参考と ならなかった	~	言えない どちらとも	~	非常に参考 となった			
全学テーマ 別評価	テーマの実施・改善のための体制	6	13	66	235	100	7	427	425
	実施方法や実施内容	2	11	35	158	71	9	286	
分野別 教育評価	教育の実施・改善のための体制	1	1	11	44	16	2	75	65
	教育内容・方法や学習支援の方法	1		10	47	19	1	78	
分野別 研究評価	研究実施・支援・改善のための体制	1	2	5	45	21	4	78	62
	研究活動のための諸施策、諸機能			8	36	11		55	
	研究内容		1	2	16	6		25	
総計		11	28	137	581	244	23	1,024	552

教育評価と研究評価にはともに「総合科学」を含む。

(出所) 筆者作成

設定をしたことなどもこの結果に影響を与えていると考えられる。

問題点としては、自己評価に携わった担当者間では、評価の目的や評価に基づく改善等の枠組みを十分に理解しているが、それが少数の人員に限られており、評価に携わった者が有する情報や経験が他の教職員には十分伝わっていないという状況にあることが各調査から明らかとなった。組織としてのさらなる質の向上のためには、情報や経験が組織の中で広く伝わるような体制の確立が不可欠であると考えられる。

(4) 大学の諸活動の社会への説明(アカウンタビリティ)に関する成果(社会で発現する「成果」)

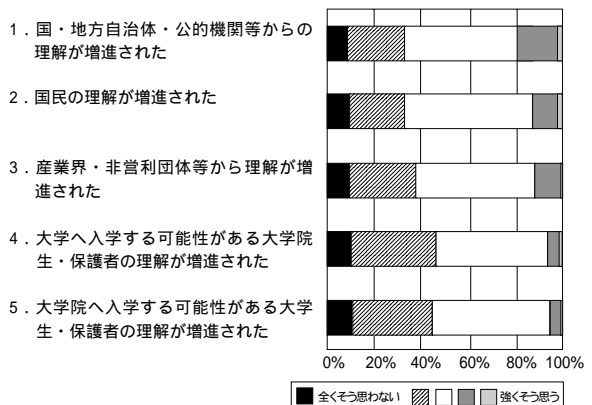
試行的評価の2つ目の目的である「大学の諸活動の社会への説明」の実現のためには、評価自体が「社会が大学の状況を把握できる評価」であることが求められた。機構の評価結果によって社会の理解が増進されたかについて質問した結果が図5である。

回答の平均値は2.5~2.8と他の質問項目に比して著しく低く、全体の3~4割が否定的な回答であった。機構では評価結果を対象機関及び設置者に

提供するとともに、マスメディアへの記者発表、印刷物の刊行、ウェブサイトの掲載等により広く社会に公表してきたが、評価委員からは、マスメディアによる報道を受けて、逆に大学の活動について誤解を生じる可能性があるというコメントすら寄せられている。

特に着目すべきは、入学希望者の活用の程度である。その他の利害関係者と比較しても低い値となっており、最も活用してもらいたい者に活用されていない状況は改善の余地がある。評価報告書

図5 関係者による大学の理解の促進



(出所) 筆者作成

を高校生やその保護者、産業界、国民一般が読むには、その内容が難しく、また分量が多すぎて、これらの人々が大学について知りたい情報が十分含まれていると言える状態にはなっていないという指摘もあった。評価担当者からも「社会に広く読まれるような分かりやすい内容と形式にすべき」との意見も多く、評価報告書の内容、あるいは評価報告書以外の公表の方法について、改善の余地が残されている。

(5) 総合的分析結果

分析結果を総合的に判断し、それをロジックダイアグラム上に表現したものが図6である。これまで分析してきた、ダイアグラムの中の「投入」から下の階層に関して、問題がなく実施されたと評価される要素は実線で囲んだままであるが、やや問題が生じたと評価される要素は点線で囲んで印を、そして問題があると評価される要素は線で囲まれずに×印を付した。

一番下の層で示されているように、目的の一つである「大学の教育研究活動等の改善」については、大学における改善にある程度の貢献が認められること、大学のマネジメントや教職員の意識改革への影響、及び大学の質の向上に関する間接的な効果も存在することなどから、まだ十分な成果の発現には時間を必要とするものの、おおむね達成できていると考えられる。その成果の発現を阻害する要因としては、ロジックダイアグラムに沿った分析からは、大学の中での評価経験の蓄積、評価の趣旨・制度等の浸透が挙げられる。前述のように、評価に携わる人間が一部にとどまり、またそこから情報共有が進まないという状況は設計時には想定されていなかった。また、大学からは、各大学の個性の促進についてより一層考慮することが望まれており、評価側にも改善すべき点が多く残されている。

一方、もう一つの目的である「広く国民の理解と支持が得られるよう支援・促進」については、利害関係者が評価結果を活用したという事例が十分には確認できず、大学評価の仕組みや評価結果に関する社会全般の理解度や活用のされ方という点での課題が多く、達成されていない。この問題

は、ロジックダイアグラム上では、評価結果の公表方法や、マスメディアによる報道がネックとして考えられる。これについては、そもそも「社会は大学に関する情報を欲している」という前提が形成されていたことも1つの原因と考えられる。つまり、社会は大学のどんな情報でも良いというのではなく、わかりやすい情報、必要とする情報を求めていると言えるが、現在のニーズからそれをランキングという形に求めるのはあまりに安易である。第三者評価機関でなくては提供し得ない、しかも利害関係者が真に必要とする情報について、高い質を保ち、必要な形で提供することが求められている。

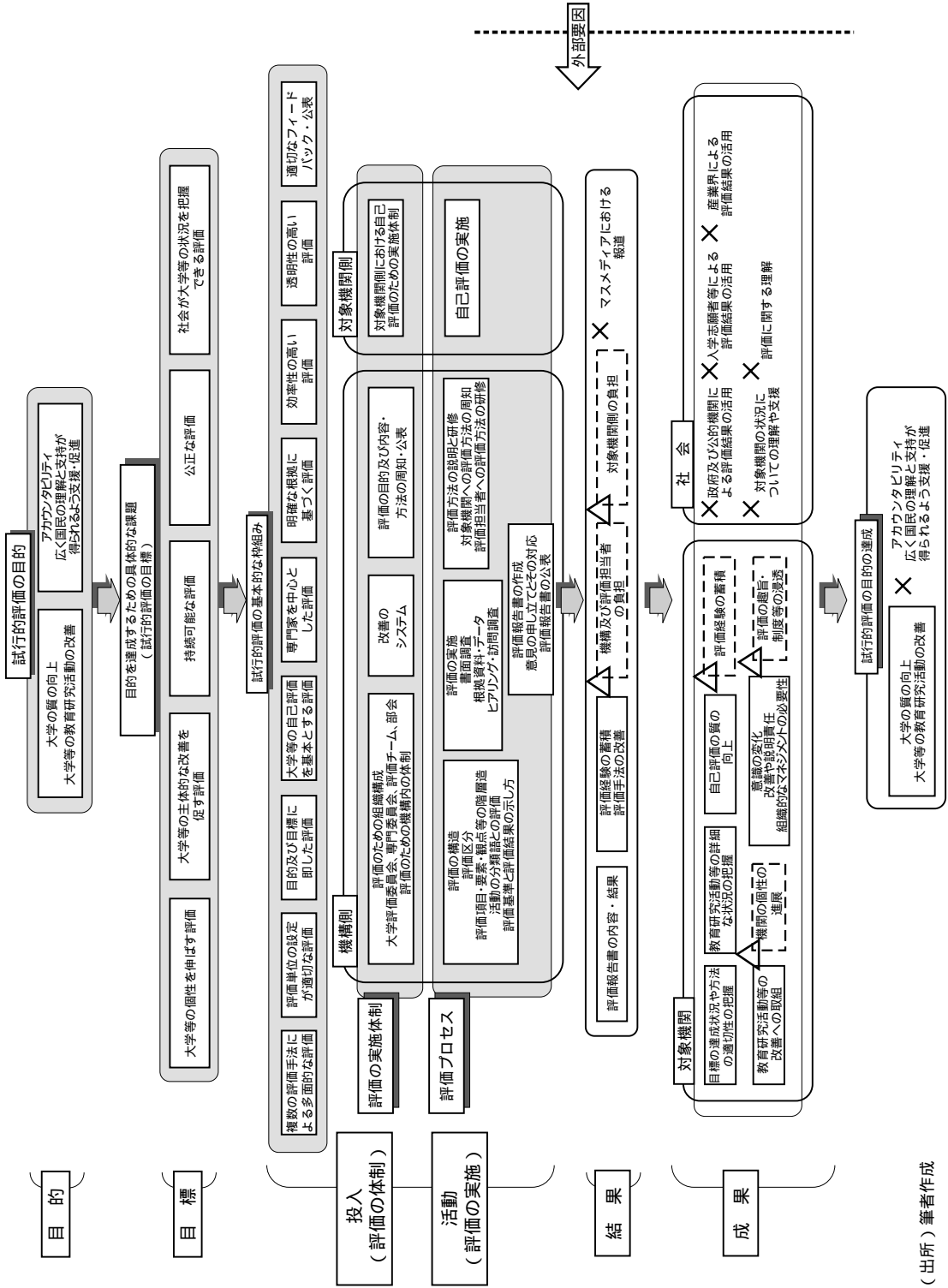
5. 作業負担の検討

本研究では当該評価の検証を行うためのロジックダイアグラムを設定する際に、あえて投入としての資源を考慮せず、作業負担という形でアウトプットに配置して分析を行った。これは、機構による試行的大学評価が今後の大学評価の基盤を作るものであるために、投入資源は所与のものとしてほとんど無視できたためである。しかし、実際には資源は無限ではなく、評価側のみならず、評価を受ける側の資源についても検討を行わなければならない。作業負担が大きいという結果は、つまりは当初の投入資源の見積もり、あるいは評価設計自体に無理があったということである。

評価に携わる者は、評価を受ける側も行う側も、社会的に見ても能力が高く、評価に費やした時間を他の業務に振り向ければどの程度の高い生産性があるのかについても本来は検討すべきであろう。大学評価は、その生産性以上に効果がある事業であることが求められるし、また評価の規模が大きくなればなるほど、評価に費やされる資源も増大する。そこで、ここでは今後の大学評価設計の参考として、評価担当者が評価に要した時間を算出し、検討を行う。

評価担当者は、2年間の評価期間中に、書面調査等の評価作業に携わるだけでなく、各種の会議に出席し、またヒアリングまたは訪問調査に参加する。これらの評価活動に要した時間には、2001

図6 試行的大学評価のロジックダイアグラム



(出所) 筆者作成

年度から2002年度にかけての2サイクル目に行った評価の中で、全学テーマの「研究活動面における社会との連携及び協力」、分野別教育評価の「工学系」、分野別研究評価の「教育系」および「法学系」のデータが利用可能であることから、それらを使用した。評価担当者による質問紙への回答であることから、あくまでも推測値である。

表2の結果を見ると、評価担当者がどの評価種別においても平均しておよそ150時間以上を評価に費やしていることがわかる。日数に換算するとおよそ1ヶ月であり、評価の主査、副主査を担当している場合にはもっと多くの時間を費やしていることを考えると、負担は極めて大きいと言わざるをえない。その主な原因は、研究成果の査読に要した時間である。試行的評価に関しては各研究者の研究成果を一定数出してもらい、それらをすべてピア（当該分野を専門とする他の研究者）が評価したことから、このような結果となっている。

対象機関あたりに要した時間に換算すると、全学テーマ別評価では対象機関の数が多いために91

時間と一番少なく、そして分野別研究評価（教育系）がおよそ3,122時間という結果になった。評価担当者が評価に費やした時間だけでこれだけの大きな負担であることから、評価対象機関の負担も加えればより多くの資源が用いられたこととなる。

これらの値からは、例えば分野別研究評価を受けた法学部では、少なくとも1,865時間の労力に相当する効果をこの評価から得られなければ、社会的に見て効果が生じていないこととなる。しかし、たとえば評価の効果が長期間持続すれば、その値は割り引かれるはずである。大学が定期的に受けることが義務づけられている認証評価は、その期間が7年間であり、そして国立大学法人が受ける国立大学法人評価の中期目標期間は6年間である。ここで単純に、対象機関あたり平均総作業時間を6年間で割れば、表2の全学テーマ（研究連携）分野別教育（工学系）分野別研究（教育系）同（法学系）の順に、それぞれ15.1時間、168時間、520時間、311時間となり、それだけの費やされた時間を1年間で取りかえさなければならない

表2 評価担当者の総機会費用

費目分類		評価種別	全学テーマ (研究連携)	分野別教育 (工学系)	分野別研究 (教育系)	分野別研究 (法学系)
(固定) 会議・打合せ			1,462	2,177	1,714	1,605
機関別 チーム	(変動) 会議・打合せ		1,915	792	525	503
	(変動) 評価作業		4,246	1,488	1,728	1,632
研究分野 別部会	(変動) 会議・打合せ				2,248	1,720
	(変動) 評価作業(査読)				12,008	5,348
(変動) ヒアリング・訪問調査			2,644	1,608	510	381
合計			10,266	6,065	18,732	11,189
評価者あたり平均作業時間			168時間	202時間	226時間	145時間
対象機関あたり平均総作業時間			91時間	1,011時間	3,122時間	1,865時間

((断りがない限り) 単位: 人・時間)

(注) 費目分類は次のように対応する。なお、(固定)(変動)とは、評価規模による増減の有無を表す。

(固定) 会議・打合せ: 専門委員会、主査打合せ等

(変動) 会議・打合せ: チーム・部会会議、説明会・研修会等

(変動) 評価作業: 書面調査、査読作業、判定等

機関別チームとは評価対象機関別に構成されたグループを、研究分野別部会とは分野(教育学、法学等)の中さらに細かい研究分野別に構成されたグループを示す。

会議・打合せに関しては、本務からの移動時間を含めていることから、実質的な作業時間は表中の値の1/3程度である。

評価作業は質問紙調査を元にした平均値で算出しており、中央値で算出すると3/4程度である。

(出所) 筆者作成

ことになる。

ある程度予測可能な範囲内とも考えられるが、研究評価は大きな負担を評価者に強いてしまったこととなる。教育評価と比較しても、研究評価の特に査読に相当する作業は軽減される必要がある。

6. まとめと提言

機構の試行的大学評価について詳細に評価を行った結果、目的の一つである大学の教育研究活動等の改善については、おおむね達成できていると評価された。一方、もう一つの目的であるアカウンタビリティの促進については、課題が未だ多く、改善の余地があると評価された。

大学の質の向上については、大学内で一部の評価担当者に作業が集中する弊害と、評価担当者の経験と情報の蓄積を組織全体に浸透させる段階が機能していないことが示された。機構としてはこのような評価に関する知識、経験の波及について、大学により積極的に働きかけていく必要がある。

また、社会からの理解については、機構や大学と、社会とを結ぶ重要な役割を担うマスメディアに重要な役割を担ってもらわなければならないことが、ロジックダイアグラム上からも明らかであり、マスメディアに対して評価の趣旨を十分に理解してもらうよう努力することが必要である。また、一方で、社会に対してそれとは別のロジック（因果関係）を用意すること、つまり別のチャンネルを用意することも重要であろう。インタビュー調査では、大学自らが情報を積極的に発信している大学は、新聞等の報道をさほど問題視していなかった。このことから、大学自らの努力で社会の理解と指示を得られるようにすることもまた必要であると考えられる。

機構の新しい大学評価制度においても、社会への因果関係をいかに成立させるかという点が懸案事項であることに変わりはなく、広報活動の充実等によって大学や社会に積極的な働きかけを行い、また評価システムを改善していくことで対応する必要がある。

機構の評価は大学の質の改善とアカウンタビリティを目的としつつも、機構の役割はそれらの支

援でしかなく、機構は評価結果を公表することはできても、そこから先の大学の質の改善とアカウンタビリティの促進は大学の手任せに委ねられ、さらに最終的な評価結果の活用は利害関係者に委ねられる。すなわち、ロジックダイアグラムの因果関係の先の方、プログラムの実施者の手が届くか届かないかという先に、成果の成就を決めるほとんどの因果関係が存在するという状況にある。

認証評価の制度の中で、多くの認証評価機関は大学の連合体という背景を有する一方で、機構は独立行政法人という設置形態である。このことが、大学の自律性との関係で、余計に協力関係を構築しづらい、手を出そうにも出せない枷となっている。

機構は、大学評価に関してはあくまでも第三者として、評価対象機関と利害関係者にどのような情報を提供していくべきか、またどのような支援を行っていくべきか、誰に対しても直接的関与はしないにせよ、評価を通じて間接的に質的向上を適切に支援していくという、難しい立ち振る舞いが要求されている。作業負担や費用効果等の評価の実施可能性とともに、大学評価の成果の発現までを見越した評価システムの設計と、因果関係が確実に成立し、成果が発現するような、すべての評価の利害関係者等への間接的支援が必要である。

謝辞

本研究は意見照会並びに各種調査に応じていただいた評価対象機関、評価者、関係団体の方々の回答を基に分析を行ったものである。ここに厚く御礼申し上げます。特に意見集計については、大学評価・学位授与機構の川口昭彦教授、芳鐘冬樹助手、野澤孝之助手にご協力いただき、また有意義なコメントをいただいた。ここに厚く御礼申し上げます。

参考文献

- Bickman, L. (1987). The Functions of Program Theory. " In L. Bickman (ed.), *Using Program Theory in Evaluation*. New Directions for Program Evaluation, 33, San Francisco, Jossey-Bass.

- 大学審議会（1998）「21世紀の大学像と今後の改革方策について - 競争的環境の中で個性が輝く大学 - 」
- 林隆之・齋藤貴浩・米澤彰純・川口昭彦（2003）「大学評価システムのメタ評価：大学評価・学位授与機構による研究評価の運営レビュー」、『研究・技術計画学会第18回年次学術大会講演論文集』、510-513
- 林隆之（2002）「欧州における研究評価の教育コースの現状」、『大学評価』、第2号、大学評価・学位授与機構、35-53
- 国際協力機構外部有識者事業評価委員会 / 2次評価作業部会（2006）「2005年度2次評価結果」、第1章、「外部有識者事業評価委員会による2次評価」、第4部、『事業評価年次報告書2005』、国際協力機構、114-140
- 国立大学協会（2002）『平成12年度着手の大学評価に対する意見について』（http://www.janu.jp/active/txt5/h14_7_2_a.html）（2006.11.1取得）
- 国立大学協会（2003a）『平成13年度着手の大学評価に関する申し入れ』（http://www.janu.jp/active/txt5/h15_6_6.html）（2006.11.1取得）
- 国立大学協会（2003b）『「大学評価・学位授与機構の評価事業の今後の在り方について（中間まとめ）」について（意見）』（http://www.janu.jp/active/txt5/h15_9_25.html）（2006.11.1取得）
- McLaughlin, J. A., and G. B. Jordan（1999）. Logic models: A tool for telling your program's performance story. *Evaluation and Program Planning*, 22, 65-72.
- 大川夏子（2002）「プロジェクト評価のメタ評価 - 人口・家族計画プロジェクトを例として - 」、『日本評価研究』、2（2）、91-100
- 齋藤貴浩（2004）「大学評価・学位授与機構による大学評価の試行に対する意見集計結果 - 全学テーマを対象として - 」、『日本評価学会春季第1回全国大会発表要旨集録』、167-174
- 齋藤貴浩（2005）「評価の評価：大学評価・学位授与機構による試行的大学評価の検証方法とその結果」、『日本評価学会第6回全国大会発表要旨集録』、257-260
- Savaya, R and Waysman, M.（2005）“The Logic Model: A Tool for Incorporating Theory in Development and Evaluation of Programs.” *Administration in Social Work*, 29, 2, 85-103.
- 試行的評価に関する検証委員会（2004）『大学評価・学位授与機構が平成12年度から平成15年度までに実施した試行的評価に関する検証について - 試行的評価に関する検証結果報告書 - 』、大学評価・学位授与機構
- Wholey, J. S.（1983）. *Evaluation and Effective Public Management*. Boston, Little, Brown.
- Wholey, J. S.（1987）. Evaluability Assessment: Developing Program Theory. In L. Bickman (ed.), *Using Program Theory in Evaluation*. New Directions for Program Evaluation, 33, San Francisco, Jossey-Bass.
- W.K. Kellogg Foundation（1998）. *The W.K. Foundation Evaluation Handbook*. Battle Creek, W.K. Kellogg Foundation. (<http://www.wkcf.org/Pubs/Tools/Evaluation/Pub770.pdf>)（2006.11.2取得）
- W.K. Kellogg Foundation（2004）. *The W.K. Foundation Logic Model Development Guide*. Battle Creek, W.K. Kellogg Foundation. (<http://www.wkcf.org/Pubs/Tools/Evaluation/Pub3669.pdf>)（2006.11.2取得）
- （2007.2.19受理）

**For Further Improvement of Universities:
Evaluation of the Evaluation Program
Implemented by National Institution for Academic Degrees and University Evaluation
in Japan**

Takahiro Saito Takayuki Hayashi

Faculty of University Evaluation and Research
National Institution for Academic Degrees and University Evaluation
tksaito@niad.ac.jp

Abstract

From academic years 2000 to 2003, the National Institution for Academic Degrees and University Evaluation (NIAD-UE) implemented evaluations of selected universities on a trial basis. Evaluation can be regarded as an intervention and subsequently the target of a program evaluation. After the NIAD-UE trial evaluation was completed, a meta-evaluation was conducted to clarify its outputs and outcomes, including successful aspects and problems, through questionnaire surveys and interviews with the evaluators and the evaluated. As a result, it was made clear that the process and the result of the evaluation had led to the improvement of universities' educational and research activities. However, many of the persons engaged in the evaluation pointed out their considerable burden. Apathy of both other university staff and the public about the evaluation also remained a problem for making an effective evaluation system. Moreover, the concept of cost effectiveness should be introduced to the new settings of university evaluations, because the evaluation cost was worse especially for research assessment.

Keywords

university evaluation, meta-evaluation, logic diagram (logic model),
cost effectiveness, self-evaluation