

○林 隆之, 齊藤貴浩, 米澤彰純, 川口昭彦 (大学評価・学位授与機構)

1. はじめに

1990年代後半以降、研究開発の領域においても、様々な評価制度が急速に制度化されてきた。研究評価の大綱の指針以降の各省庁や研究所における評価、大学の第三者評価、独立行政法人の評価、研究施策も含めた政策評価法の施行などである。だが、急速な制度化の中で実際には試行錯誤が続いている。どのような評価システムを設計すれば良いのか、どのような評価項目や方法が良いのか、評価者をいかに選定すれば良いのか。これらの問いへの解答のないままに暫定的に評価が進められる中で、最も懸念されることは、適当でない評価が行われ続けることにより、効果が上がるよりも前に、評価に対する疲弊ばかりが増すことである。

そもそも評価システムを改善させるためには、一つには、評価システムや方法への理解を深め、専門知識を有する人材を育成することが必要である¹。それと共に、現在行われている評価活動を検証する(メタ評価)ことにより、不断無くシステムを修正していくことが求められる。

そのため、本稿では大学評価・学位授与機構(以下、機構と略す)が2000年から試行的に開始した大学の第三者評価の検証を行う。評価システムのメタ評価は、大きく分けて、評価のプロセス(評価の枠組みの設計、実施、評価結果の生成)が適切であったかという点と、その評価によって評価対象にどのような効果が得られたのかという点の2つを考える必要がある。本稿では、この内の前者についての分析を行う(後者については、評価から数年後に判断すべきであるため、今後分析を行うことを計画している)。

本稿では、機構が評価を実施する中で行ってきたアンケート調査の結果を基にして評価システムの課題をまとめる。特に、機構の評価システムに固有の問題よりも、研究評価一般に内在する課題について整理を行うことで、含意を提供する。

2. 機構の評価システムの概要

機構では、学部ごとの研究評価、教育評価、および大学ごとのテーマ別評価の3種類を試行的に行っている。本稿ではこの内、研究評価に焦点をおく。研究評価はこれまで、理学、医学(1年目)、工学、教育学、法学(2年目)の5分野について、それぞれ6つの大学の学部あるいは附置研究所を対象に評価を行った。機構の評価のプロセスは図1のようなものである。

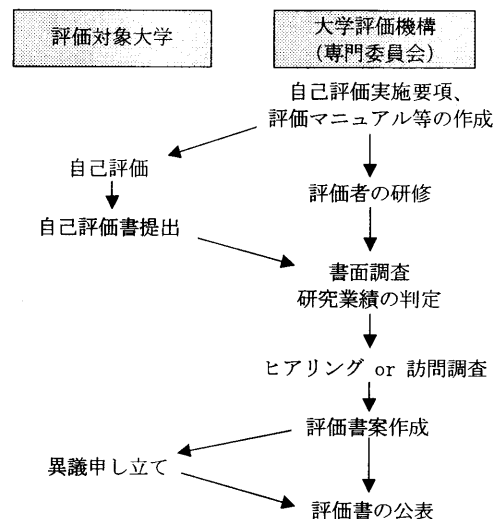


図1 大学評価機構の評価プロセス

基本的な構造は、評価対象の大学自身が自己評価を行い、それを基に機構に設置された専門委員会が評価を行う。ただし、自己評価は機構側が作成したガイドラインに沿って行われ、根拠資料を示すことが求められる。専門委員会は自己評価書の書面調査とヒアリングあるいは訪問調査により最終的な評価を行う。

評価は次のような項目ごとに行われた。

- 1) 研究体制および研究支援体制
- 2) 研究内容および水準
- 3) 研究の社会(社会・経済・文化)的効果
- 4) 諸施策および諸機能の達成状況
- 5) 研究の質の向上および改善のためのシステム

¹ 例えば評価人材の育成については、林隆之「欧州における研究評価の教育コースの現状」『大学評価』Vol. 2, pp. 37-53を参照。

表1 分析対象とした「意見照会」

No	H12-13年度の評価		回答数
①	実施要項に対する意見	関係団体	10
②	自己評価の方法への意見	対象大学	10
④	H12の大学評価への意見	専門委員・評価員	72
⑤	評価結果への異議申し立て	対象大学	9
⑥	H12の大学評価への意見	対象大学	11
H13-14年度の評価			
③	実施要項に対する意見	関係団体	6
⑦	自己評価の方法への意見	対象大学	16
⑧	H12-13の大学評価全体への意見	大学評価委員	8
⑩	H13の大学評価への意見	専門委員・評価員	100
⑪	評価結果への異議申し立て	対象大学	8
⑫	H13の大学評価への意見	対象大学	15
H14-15年度の評価			
⑨	実施要項に対する意見	関係団体	7

*Noは実施順。回答数は研究評価への回答のみを集計。

この内、1)、4)、5)は研究組織としての体制や方策に関する評価を行う。各項目の中にはさらに小項目(「要素」)が設定され、それぞれについて評価者は優れた点、改善点を指摘するとともに、4段階(H12年度)あるいは5段階(H13、14年度)の評価をつける。

他方、2)および3)は、過去5年間の研究の成果、およびその社会的効果を評価するものである。試行的評価では、評価対象の学部構成教員それぞれから、研究内容、研究業績リスト、研究業績そのもの5点以内を提出してもらい、それを専門分野ごとの部会によるピアレビューで4段階に「判定」し、その結果を集計して組織の評価とする方法をとった²。

3. 分析対象のアンケート

本稿のメタ評価で対象とするのは、機構が評価対象大学や評価者に行った「意見照会」(郵送によるアンケート)の結果である。機構では、評価システムを不断なく改善することを目的に、適時、意見照会を行ってきた。これまで行った意見照会は表1の通りである。

これらは、統計的分析を行うことを目的に設計されたものではない。評価システムの課題を探索的に把握し、特に、次年度の評価システムを修正することを目的とするものであった。そのため、意見紹介の様式は選択式の質問ではなく、設定された少数の質問項目ごとに自由記述してもらった形式であった。そのため、本分析では、比較的似た意見を集約して回答数を数えることで集計を行った。

4. アンケート結果

アンケートには多種のコメントが寄せられた。表2はその内の主要なものの一覧である。上述のように、研究評価は、組織の体制・方策に関するものと、研究成果や社会的効果に関するものに分けられる。そのため、アンケート結果も分けて示している。本稿では、これら集計結果の中でも、容易に次年度に修正対応が可能ではないような、研究評価の本質的な課題としてどのようなものが指摘されたかを以下にまとめる。

4. 1 組織の研究体制・方策の評価へのコメント

(1) 評価を行う目的の共有

評価を行う目的の認識が評価実施側と評価対象側との間で実際に一致しないと、評価の実施に問題を生む。

機構の行う評価の目的は、大学の改善の促進と大学活動の社会への説明にある。だが、評価を受ける大学側は、外部からの一方的な監査へと同様の対応を取りうる。コメントには、大学が良い評価結果を得るために甘い自己評価を行うことや、目的・目標を低く設定することの問題が指摘されている。さらには、研究成果だけでなく、研究組織の評価まで行うことに意味があるのか、改善目的ならば評価が改善に活かされる制度化が必要である旨が指摘されている。

(2) 組織の目的・目標の明確設定の不慣れ・困難さ

組織評価においては、その組織の目的・目標に即して評価を行うことが一般的である。これは、多種多様な対象を多様なままに評価して、改善を促進しようとするためである。しかしながら、特定の研究分野に限られる研究所ではなく、大学の学部という個人の共同体的組織では、組織の目的・目標を明示的に設定した経験が少なく、設定の困難さを指摘するコメントが多い。結果、「世界レベルの研究の実施」などの抽象的なものが書かれ、評価の実施も難しくなっている。

(3) 特定機能を評価することによる影響

評価では対象の多様な機能の中の一部に焦点を絞って評価を行う方法がとられる。しかし、他機能が同等に評価されるという補償がない場合には、評価対象となった特定の機能のみを奨励するという影響を持つ。機構の大学評価の場合には、教育や診療活動(医学部)などを軽視する影響があるとコメントされている。

(4) 評価コストと評価者のモチベーション

評価にかかる労力の大きさと作業期間の短さは、評価対象大学および評価者の双方から指摘されている。

² ただし、3年目の「総合科学」分野では、個々の教員ではなく、その組織の研究活動を幾つかの「研究プロジェクト」としてまとめて、それらについて代表的研究成果を提出してもらった方法をとっている。

特に、大学からは自己評価を行う際に根拠資料を収集・整理することの労力の大きさが指摘されており、日常的な情報収集体制の整備の必要性も指摘されている。また、評価者にとっては、質の高い評価を行おうとすればするほど作業量が増えるというジレンマを抱えている。これは、研究成果の判定作業も同様である。

(5) 評価者の研修

評価者の研修の必要性は評価の質の向上にとって必要というコメントが多い。評価対象大学にとっては評価結果の受容性を高め、評価者にとっては自己の評価作業への確信と効率性という点から求められる。

4. 2 研究成果と社会的効果の評価へのコメント

(1) 評価を行うレベル・単位の設定

本来、自己改善や説明責任を目的とする組織評価では、評価対象組織の内部に構成員とその成果を評価する体制が確立していれば、第三者評価機関がそのレベルの評価を詳細に行う必要はない。しかしながら、日本ではそのような体制が確立している大学が少ないことや、産業界などから各大学の研究水準を明確に示すことを求める声があることから、研究成果の評価のみ教員個人の判定作業を行い集計をとる方法を採用してきた。さらに、ピアレビューを多数の大学に対して効率的に行うために、各大学の組織構成とは関係なく、機構で設定した専門分野ごとに「部会」を設定して判定作業を行うシステムを形成した。

組織評価でありながら、組織構成とは異なる区分で個人の判定作業を行う設計に対して、「個人単位で行うべきでない」「個人に評価結果を返すべき」「研究は個人活動であるので個人評価で良い」など異なる意見が寄せられている。これらは評価目的と評価コストとも関係して、概念の再整理が求められる。

(2) 分野ごとに適した様式と比較可能性

多様な研究分野の評価を行うため、分野ごとに評価書の様式、判定基準、提出する業績の種類などを設定すべきという意見がある。他方で、分野固有の様式をとることにより、分野間での比較可能性が下がり、評価結果が誤解を招くという意見もある。このトレードオフのバランスをいかにとるかは、分野を超えた評価の根本課題であり、英国RAEでも認識されている。

(3) 評価の公平性と指標等による支援

評価の公平性で問題となるのは、評価者による判定基準の違いである。そのため、判定基準をより具体的に明示化することが求められており、機構でも毎年修

正を行ってきた。だが、「社会的効果」をより詳細に定義する必要など、残された課題も多い。

またピアレビューという主観的判断の質を高めるためには、より客観的な参照情報を提供することによる支援が必要となる。だが、指標にも分野間差異や誤用の問題が指摘されており、それらを把握した上での評価が必要となる³。また、評価対象をカバーできる評価者を取りそろえることの難しさも指摘されている。

(4) 短期的視野による評価の是非

研究評価に常に生じる問題の一つが、数十年スパンで判断されるべき研究の効果を5年程度の短期に評価することが求められることである。根拠資料を求める評価においては効果を示すことが困難という指摘もあり、限界を前提として、効果の顕在化の構造の評価などが必要となる。

5. 評価システムの改善に向けて

大学という組織の研究評価の方法はいまだ確立したものではない。英国RAEも見直しの最中であるし、日本でも、今後、国立大学法人評価に向けた新たなシステムの設計が必要になる。むしろ評価は一意に最適な方法を確認できるものではなく、評価方法自体の改善とともに、「評価の浸透度」に合わせた設計が必要と言える。仮に自己評価体制が確立していなければ、第三者評価機関は踏み込んだ濃密な評価を行うとともに、自己評価体制の形成を推奨することが要求される。逆に、評価対象大学が継続的にデータ収集を行っていれば、それをういた簡素な方法や、ある程度客観的な情報提供も可能となる。

アンケート結果から得られた、現時点の日本の大学の研究評価の課題は次のようにまとめられる。

- ・ 評価の目的と評価対象の理解の統一
- ・ 評価を含む大学の運営サイクル全体の設計
- ・ 自己評価の内在化と情報収集体制の整備
- ・ 評価の支援方法の改善
- ・ 評価の限界の認識とその下での評価結果の利用

謝辞

意見紹介に回答いただいた評価対象機関、および評価者の方々には厚く御礼申し上げます。また、本集計は報告者らとともに、芳鐘冬樹、野澤孝之の両氏の協力のもとで行った。

³ ピアレビューと論文数・引用数などの指標との関係については、林隆之「ビブリオメトリクスによるピアレビューの支援可能性の検討」『大学評価』Vol.1.3（印刷中）参照

表3 主なコメントと回答数

		研究体制・方策などの評価に関するコメント										回答数(H12)				回答数(H13)				H4				
		①	②	④	⑤	⑥	③	⑦	⑧	⑩	⑪	⑫	⑬	⑭	⑮	⑯	⑰	⑱						
評価の目的		評価の目的が不明。この評価で研究が活性化されるか疑問改善への連結の制度化の必要	1	3													1							
大学の目標設定		目的・目標自体の評価の必要性		1		2							1	5			1							
		研究内容についての目標は上位にいくほど抽象的になり、下位にいくほど分散的で細分的なものになる	1									1	1	1										
		制度上の制約や大学全体レベルで扱うべきことは学部で対応しにくい。目標に書きにくい	1	1	1												1	2						
他機能との関係		研究活動を組織の目的・目標との関係で評価するのは無理	2		1												2							
評価方法	評価項目・要素・観点	「要素」を固定することで大学の画一化につながらないように注意すべき。共通的なものと選択的なものを分けるべき	2	2	1		2						1						2					
	自己評価のデータ収集	根拠資料を過去に遡って集めることが困難								1	1													
	ヒアリング・訪問調査	継続的にデータを収集することが必要	1										1						2	1				
	水準付け	ヒアリングが短すぎる、訪問調査を実施すべき、訪問調査を行って良かった(工学)			5														7					
	評価報告書	水準をつける基準が不明瞭																	6					
	研修	作業量の割に、成果が乏しい																	4					
専門委員	研修	評価者の研修をもっとやるべき	7		1								1	13										
	長期的育成	同じ委員を継続して参加させるべき。専門家の育成が必要。2～3年で委員を変えるべきでない	1		1									1	3									
スケジュール		スケジュールがタイト、1年では短い	4	3								1	2	8										
作業量		根拠資料の作成の作業量が多い	2	2		2						3	1	3				2	2					
		評価者の作業量が多い。より簡素化すべき	2	4		1								1	10					1				
		研究内容と水準・社会的効果の評価に関するコメント										①	②	④	⑤	⑥	③	⑦	⑧	⑩	⑪	⑫	⑬	⑭
相対評価の是非	大学の目的・目標との関係	大学の研究の目的・目標を考慮していない、大学としての特性性を評価すべき	1	1	3			2									1							
	環境条件を考慮した評価	大学の規模、設備、研究費、大学院生の数を考慮していない							3	3							2							
評価を行う単位		他機能(教育、診療など)による制約、他の制約(政府、大学全体の責任事項)を考慮していない	1		1			1												1				
		個人ごとの判定を行う必要はない。組織別に行うべき。個人の評価(判定)は大学自身がやるべきこと	2		4		1	1					1											
		個人→部会→研究科という統合方法により、学科が抜け落ちることの問題(学科レベルこそが意志決定には重要)	1				1					1												
		個人に結果を返さないため、個人の改善にならない		1	1	1	1										2			2				
	大学内および外との共同研究(共著)が多い場合に、個人単位で評価することの問題	1		1		2										2			1					
部会の分野設定		対象大学にあわせた分野設定の必要、学際的の分野や特定分野に特化している大学の扱いの問題	3		1	2	2					3	1	3						1				
評価の方法	対象期間	5年では短すぎる															1		9					
	対象者	若手を同一に評価すべきでない															2	1	3	7				
	自己評価、申請書	研究活動判定表の書き方が大学・個人で異なる。控えめな自己申請書類が低く評価される可能性															3	1		1				
	語の定義	新規性、発展性、創造性、独創性の定義が不明															5		1	4				
	分野ごとのフォーマット	分野にあったフォーマットや具体的例示が必要															2	3	1	4				
		「国際的視点」の定義が不明、「国際的」が適さない分野がある	2	1									1	1					5	1				
	社会的貢献・効果	「社会的貢献」「社会的効果」の定義や判断基準が不明。															1	15		2				
		各種の社会的効果の定義も明確にすべき																		1	2			
		社会的側面があつての研究(学術面に重点がおかれすぎ)																		3				
	社会的効果の評価可能性	効果・インパクトの種類例示の不十分性	1																1	3	1			
	効果・インパクト顕示まで5年では短く、根拠資料がでない。因果関係が不明で明示的な根拠資料が示せない場合も																		2	6				
指標	指標の必要性																		9	2				
	論文数・引用数は使うべきでない	1																		1				
判定基準		判定基準の具体的明示化が必要	1	6	13	1	4												5	1				
		評価員間での見解の相違がある																	2	1				
		部会間での判定基準が不統一。定義が曖昧																	1	4				
		そもそも部会内部でも統一した基準を作ることは困難																	1	6				
		評価対象となる業績の種類、重み付けを明確にすべき																	5	2				
他機能との関係、評価のインパクト		個人の教育活動、診療活動、社会貢献活動、機器・安全管理などが軽視されることを助長する	2		1														2	8				
			2	2	4															3				
評価員	評価員の人数・構成	評価員を増やすべき																	5	9				
	研修	評価員では判断しきれない研究があつた																	4	11				
スケジュール		評価員の研修の必要																5	1	6				
作業量		判定作業のスケジュールがきつい																	3	5				
		自己評価にかかる作業量が多い	1	1															1	1				
		判定にかかる時間が大きい																	3	15				
																				1				