

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

Model of Institutional Quality Enhancement Promoted by “Certified Evaluation and Accreditation”
: Comparative Analysis among Different Types of Higher Education Institutions

金 性希, 林 隆之, 齊藤 貴浩
KIM SoungHee, HAYASHI Takayuki, SAITO Takahiro

1. はじめに	21
2. 各種の認証評価の概要	22
3. 検証アンケート調査の概要	24
4. 学校種別による回答の比較	24
4.1 評価基準および観点	25
4.2 評価の方法および内容	25
4.3 評価の作業量	27
4.4 評価に関する説明会・研修会	28
4.5 評価結果（評価報告書）について	28
4.6 評価を受けたことによる効果・影響について	29
4.7 まとめ	31
5. 構造分析	32
5.1 探索的因子分析	32
5.2 因果モデルの検証	32
5.3 対象校間の比較（多母集団同時分析による因果関係の比較）	35
6. 議論	36
ABSTRACT	42

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

金 性希*, 林 隆之**, 齊藤 貴浩**

要 旨

本研究では、大学評価・学位授与機構が実施した大学・短期大学、高等専門学校、法科大学院を対象とする認証評価について、評価対象校へのアンケート結果を基に、対象校の種別や評価システムの性質によって評価の適切性の認識や改善効果が生じる構造に差違があるかを検討した。対象校間の比較の結果では、高等専門学校の認証評価は、対象校がいずれも工学分野であるなど特徴がほぼ均質であるために評価者との共通認識が形成され易く、評価による改善効果も高い。一方、法科大学院は、プログラム単位の適格認定という評価システムの性質から、詳細な評価基準への適合性が重視され、評価者と対象校との見解の相違が顕在化しやすいことが示された。今後、日本で分野別評価が導入される場合には、教育内容や方法の質保証の厳密性と対象校の自由度の確保のバランスが課題となることが示唆された。構造分析の結果からは、対象校での改善のためには自己評価による適切な現状把握が重要であり、同時に第三者評価の結果が外圧やインセンティブとして働くことが示された。将来的に、大学等で内部の恒常的な質保証体制が十分に構築されると、第三者評価に起因する新たな改善効果が逡減することが本分析から予想されることから、より費用対効果の高い評価方法へと変更していくことが求められる。

キーワード

認証評価, 評価の検証, メタ評価, 多母集団共分散構造分析, 評価の内部浸透段階仮説

1. はじめに¹

2004年の学校教育法の改正により、全ての高等教育機関は7年に1度、全学的な教育研究等の状況に関する第三者評価を受けることが義務づけられ、また、専門職大学院は別途5年に1度の評価を受けることが義務づけられた。この評価は、文部科学大臣の認証を受けた評価機関が行うものであり、それゆえに「認証評価」と呼ばれる。

この認証評価の制度化以前に、大学評価・学位授与機構（以下「機構」と表記する）は、1998年の中央教育審議会答申に基づき、2000年から国立大学を主な対象とした「試行的な」大学評価を

実施していた。これは3年後に本格実施に至る予定であったが、上記の法改正、ならびに国立大学の法人化という高等教育政策の変化により中止となり、機構は評価事業の一つとして新たに認証評価を実施することとなった。機構による認証評価は2005年度から開始し、2007年度末までの3年間で大学52校、短期大学5校、高等専門学校56校、法科大学院29組織（予備評価および本評価の延べ数）の評価を終えた。

日本では大学の第三者評価は新しい営みである。そのため、機構では「進化する評価システム」を標榜し、実施した評価についてその適切性や有効性の検証（メタ評価）を行うことで、評価方法を

* 東京工業大学大学院 博士課程

** 大学評価・学位授与機構 評価研究部 准教授

¹ 本稿は、検証アンケート結果の分析を基にした筆者らの個人的研究の成果であり、大学評価・学位授与機構の公式な見解を示すものではない。

継続的に修正し改善することを目指してきた。試行的評価については、評価実施中ならびに終了後にアンケート調査やヒアリング調査を行い検証を実施した(大学評価・学位授与機構 2004, 林, 齊藤 2005, 齊藤, 林 2007)。また、認証評価へ移行した後も、毎年の評価終了後に、評価を受けた機関・組織(以下「対象校」と表記する)、および評価者の双方にアンケート調査を行い、評価の継続的改善へと繋げてきた(大学評価・学位授与機構 2007a, 2007b, 2009)。

このように機構では毎年の検証を行うことで、評価方法の具体的な課題を検討し、順次それらを修正する体制をとっているが、一方で、個々の認証評価を超える課題の検証は残ったままである。それは、認証評価という制度全体に絡む課題であり、重要なものとして以下の2つの点を挙げることができる。

一つは、現在、複数の認証評価機関により行われている評価の間で、いかなる相違があり、それによっていかなる効果・影響が対象校ならびに日本の高等教育全体に生じているかを明らかにすることである。認証評価の導入を打ち出した2002年(平成14年)の中央教育審議会の答申『大学の質の保証に係る新たなシステムの構築について』では、「大学の理念や特色は多様であるため、各々の評価機関が個性輝く大学づくりを推進する評価の在り方に配慮するとともに、様々な第三者評価機関がそれぞれの特質を生かして評価を実施することにより、大学がその活動に応じて多元的に評価を受けられるようにすることが重要である」として、複数の認証評価機関が評価を行うシステムを構想し、実際に、機関別認証評価、専門職大学院認証評価ともに、現時点で複数の機関が存在している。それらの中で評価基準の内容や水準、評価方法の透明さや厳正さ、評価によりもたらされる影響・効果などを比較し、各評価の特徴を明らかにするとともに、評価を通じた高等教育の質保証体制が国全体として適正に機能しているかを検討する必要がある。例えば、羽田(2007)は3つの大学機関別認証評価機関の評価基準を比較し、共通性は多いが評価項目の重点には相違があると述べている。このような評価基準の構成比較に加えて、評価が質保証のために適切な時間や労力をかけて厳正に行われているか、対象校でどのような

改善が促進されたか、国際的通用性が担保されているかなど、運用面や効果を含めた検討を今後行う必要がある。

もう一つは、対象が異なる認証評価の間の差異である。現時点では、大学、短期大学、高等専門学校といった機関単位の評価と、法科大学院や経営大学院をはじめとする各種の専門職大学院を単位とする評価が行われている。さらに中央教育審議会では、2005年(平成17年)の答申『新時代の大学院教育』において「機関別評価」に加えて、大学院教育の専門性に沿った「専門分野別評価」を導入していくことが適当であると述べ、2008年(平成20年)の答申『学士課程教育の構築に向けて』では「第三者評価制度の見直しに当たっては、分野別の評価をどのように進めていくかが重要な課題となる」と述べるなど、専門職大学院以外の各学位課程においても分野別評価の導入の必要性を指摘している。今後、分野別評価の導入が本格的に検討されるとすれば、機関別評価と分野別評価の相補性や整合性をいかに構築するのか、評価制度の浸透とともにそれらをどのように変化させるかが問題となろう。そのため、少なくとも現時点までに行われた対象の異なる評価において、いかなる差が生じているのかについての事実を積み上げることが、今後の評価設計のために求められる。

本稿はこの二つ目の点に資するものである。本稿は、機構が行った認証評価の検証アンケートを用い、対象の異なる認証評価において、その適切性や効果の認識がいかに異なるか、さらには、機関・組織において改善が促進される構造がいかに異なるかについて、統計的に明らかにする。具体的には、適切性や効果の認識に関する対象校による差異を分散分析で、そして教育研究の改善にいたる包括的構造、および対象校の学校種別間の比較を共分散構造分析で明らかにする。これにより、今後の機関評価および分野別評価の構築・修正のあり方への含意を得ることを目的とする。

2. 各種の認証評価の概要

具体的な分析を行う前に、まず、対象の異なる認証評価の評価方法の違いをまとめておく。

機構では、認証評価を必要とするすべての学校種(大学、短期大学、高等専門学校)を対象に機

関別認証評価を行っており、またそれに加えて専門職大学院のうち法科大学院を対象とした専門分野別認証評価を行っている（実施数については後述の表1を参照）。

評価の手続きに関してはこれらでほぼ共通である。すなわち、いずれの評価も、機構が定める評価基準に基づいて評価対象となる機関や組織が自己評価を行う。機構に設置された評価委員会は、この自己評価の結果（自己評価書）に基づき書面調査を行い、訪問調査において大学関係者（責任者）や教員・学生との面談等により不明点の確認を行う。そして、両調査の結果をもとに評価結果案を作成し、大学からの意見申立ての機会を経て、最終的な評価結果を確定する。評価結果（評価基準を満たしているか否か、優れた点、改善すべき点、特色ある点、およびそれらの評価結果を導いた理由）は、報告書として、当該機関および設置者に通知するとともに、広く社会に公開する。

このように、各評価で手続き的には共通性があるが、評価の法的位置づけ、基準、評価対象の単位という点で、差異が存在する。

認証評価の法的位置づけについては、学校教育法第109条は大学²が機関別および専門職大学院の評価を受けることのみを求めており、米国のアクレディテーションのように大学等の適格認定を行うことを求めている。つまり、機構の行う機関別認証評価は評価基準を「満たしているか」あるいは「満たしていないか」の判断を行うが、大学、短期大学、高等専門学校に関しては機構が定めた評価結果の示し方に過ぎず、法律上の要請ではない。しかし、法科大学院の場合に限り、「法科大学院の教育と司法試験等との連携等に関する法律」（以下「連携等に関する法律」と表記する）第5条において「認証評価では当該法科大学院の教育研究活動の状況が法科大学院認証評価基準に適合しているか否かの認定をしなければならない」として適合性の判定が要請されている³。したがって同じ認証評価でも、法科大学院とそれ以外の学

校種では、法科大学院が法的に適格認定を求められているという意味で、適合可否という評価結果が有する重要性も異なる。

評価基準については、それぞれの認証評価機関が定める。しかし、学校教育法ではその基準を適用するに際して必要な細目は文部科学大臣が定めるとされており、「学校教育法第百十条第二項に規定する基準を適用するに際して必要な細目を定める省令」で次のように規定されている。大学（含短期大学）の評価基準に必要とされる事項は、教育研究上の基本となる組織、教員組織、教育課程、施設および設備、事務組織、財務、その他教育研究活動等に関することとされている。高等専門学校もこれが準用されている。一方、専門職大学院の認証評価は教育プログラムの評価であるため、教員組織、教育課程、施設および設備、その他教育研究活動に関することが事項として要求されており、機関を単位とする評価からは一部が免除された形となっている。しかし、専門職大学院の中でも法科大学院だけは、その修了が司法試験と密接な関係を持つ設計がなされたために例外であり、より詳細な評価基準、例えば入学者の多様性の確保、教育課程の編成、同時に授業を行う学生の数の設定、履修登録単位数の上限設定などの評価が要求されている⁴。このため、機構の定める法科大学院の評価基準でも、大学、短期大学、高等専門学校に比べてより詳細な評価基準が策定されている。

評価対象の単位の点では、法科大学院認証評価は専門職学位プログラムを単位とする分野別評価の一つであり、その他は機関別認証評価である。ただし、実際には高等専門学校の認証評価にも特殊性が存在する。それはほとんどの高等専門学校が小規模で、その教育課程のほとんどが工学系の学問領域にあるため、機関別の評価といえども分野別評価に近い様相を呈するということである。評価基準も教育内容や教育方法等については工学分野に限定した詳細な設定が可能であり、PBL

² 大学には短期大学を含む。また、第123条において高等専門学校にも準用されている。

³ 後述する「学校教育法第百十条第二項に規定する基準を適用するに際して必要な細目を定める省令」でも、評価方法が「連携等に関する法律」に規定される認定を適確に行うに足りるものであることを求めている。

⁴ 「連携等に関する法律」では、第二条で「法曹養成の基本理念」を定め、第五条で文部科学大臣は学校教育法第百十条第三項に規定する細目を定めるときは、大学評価基準の内容が法曹養成の基本理念を踏まえたものとなるように意を用いなければならないとしている。

表1 年度別調査対象数

対象	平成17年度		平成18年度		平成19年度		合計	
	対象校	回答 (回答率)	対象校	回答 (回答率)	対象校	回答 (回答率)	対象校	回答 (回答率)
大学・短大	6 (2)	5 (83.3%)	11 (1)	11 (100%)	40 (2)	38 (95.0%)	57 (5)	54 (94.7%)
高専	18	18 (100%)	18	18 (100%)	20	20 (100%)	56	56 (100%)
法科大学院	4	4 (100%)	13	13 (100%)	12	11 (91.7%)	29	28 (96.6%)
合計	28	27 (96.4%)	42	42 (100%)	72	69 (95.8%)	142	138 (97.2%)

※以下の全ての図表において、短期大学を短大、高等専門学校を高専と略す。

※ただし、大学・短期大学の評価対象校数のうち、括弧内は短期大学の内数。

(プロジェクトベース学習) やインターンシップ等の工学教育のニーズを観点の中に取り入れており、また、後期中等教育に相当する教育課程も対象となることから一般教育との関係や生徒指導を観点に含むなど、大学とは異なる評価基準の設計をしている。また評価委員会のメンバーも必然的に工学系の評価者が多くを占めるものとなっている。

これらのことから、機構の行う認証評価は、大学・短期大学、高等専門学校、そして法科大学院と、それぞれに特徴を有したものとなっている。

3. 検証アンケート調査の概要

認証評価の検証アンケート調査は、平成17年度から19年度の3年間にわたり、認証評価を受けたすべての対象校および評価者を対象として、各評価終了直後に実施した。郵送による記名式のアンケート調査である。このうち、本稿では教育研究活動が評価により改善する構造を明らかにするため、対象校からの回答について分析を行う。

表1に評価の対象校数、およびそのうちの検証アンケートへの回答数を示す。全体で回答数は138校であり、97.2%という高い回答率である。評価開始からの3年の間に、評価方法や評価の観点等についてはいくらか修正がされているが、評価基準の抜本的な修正等はなされていないため、基本的に同一の評価と考えることができるものとして、本分析では3年分の回答をまとめたデータを分析に用いる。

検証アンケートの内容は年度・対象間で若干の違いはあるものの、多くの質問項目は共通である。アンケートは表2で示す9つのカテゴリからなり、試行的評価の検証で用いたものに準じて評価実施の目的・目標、評価方法、評価結果、評価に

表2 対象校に対するアンケート項目

問1	評価基準および観点について
問2	評価の方法および内容について
2-1	自己評価について
2-2	訪問調査等について
2-3	意見の申立てについて
問3	評価の作業量、スケジュールなどについて
3-1	評価に費やした作業量および機構が設定した作業期間について
3-2	評価作業に費やした労力について
3-3	評価のスケジュールについて
問4	説明会・研修会などについて
問5	評価結果(評価報告書)について
問6	評価を受けたことによる効果・影響について
6-1	自己評価を行ったことによる効果・影響
6-2	機構の評価結果による効果・影響
問7	評価結果の活用について(記述式・多肢選択式)
問8	評価の実施体制について(記述式)
問9	その他(記述式)

より生じる効果・影響を概念整理したロジックダイアグラムを作成し、各項目に対応する質問項目を作成した(詳細は、齊藤、林2007を参照)。すなわち、評価を実施する目的に対する評価設計の適切性、評価方法の運用上の適切性、評価結果の妥当性、評価により生じる正負の効果・影響について、選択式回答(ほとんどが5段階尺度)で回答を求め、さらに一部の事項を自由記述で補完した。

4. 学校種別による回答の比較

以下では、上記の質問カテゴリごとに、対象校を大学・短期大学、高等専門学校、法科大学院の3グループにわけ、アンケート回答結果を説明する。さらに、分散分析によってそれぞれのグループの回答結果の比較を行う。統計的に有意水

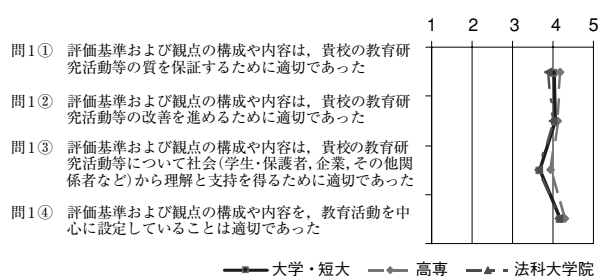
準10%までの差が見られた項目については、その後の多重比較を行い⁵、どの対象校種の間で差があるかを表で示す。分析に用いるのは、全ての対象・年度に共通する5段階の選択式回答である。ほとんどの質問では、1が「全くそう思わない」、5が「強くそう思う」となっている。3の「どちらとも言えない」が中立回答であるが、この種のアンケートでは肯定的な回答がなされやすい傾向があるため、平均値3.5程度を中立的な回答として解釈する。

4.1 評価基準および観点

最初の質問は、機構の認証評価の3つの目的である「教育研究活動等の質の保証」「教育研究活動等の改善」「教育研究活動等について社会から理解と支持を得る」のそれぞれに対して、機構の評価基準および観点の構成や内容が適切であったかを問うている。

図1で示すように、対象校の種別を問わず、3つの目的のうち、「質の保証」「改善」は5段階中4前後と適切性が高く認識されている。残りの「社会からの理解と支持」という目的への適切度についても3.5を上回る肯定的な回答ではあるが、他の2つの目的と比べれば低い結果となっている。この傾向は試行的評価の検証でも指摘されていたものである。質保証や改善を目指して作成された評価基準では、例えば教員のFD（ファカルティ・ディベロップメント）活動や内部質保証体制などの内向きの観点に重点がおかれやすい。それらは最終的には大学の質に結びつく、間接的な質の指標ではあるものの、必ずしも学生や企業などの大学外部の関係者が直接的に関心を抱くものとはなりにくいことから、このような回答結果となっていると考えられる。

対象校3グループ間の比較では、全ての項目において高等専門学校の回答結果が高く、大学・短期大学、法科大学院の順となった。表3に示す分散分析とその後の多重比較の結果では、「質保証」において高等専門学校と法科大学院で有意差が認められた。法科大学院においても回答平均値は3.89と高かったことを考えれば、高等専門学校が



(注) 尺度 1: まったくそう思わない ~ 3: どちらとも言えない ~ 5: 強くそう思う

図1 評価基準および観点についての項目ごとの平均値

表3 評価基準および観点についての分散分析の結果

項目	F 値	グループ1		グループ2		グループ3		多重比較		
		大学・短大		高専		法科大学院		グループのペア		
		平均	N	平均	N	平均	N	1と2	1と3	2と3
問1① 教育研究活動等の質保証に適切	3.091*	4.04	54	4.18	56	3.89	28			◇
問1③ 社会からの理解・支持を得るに適切	2.510 [†]	3.69	54	3.95	56	3.64	28			

* p<.05, [†] p<.10

多重比較の例では、有意差（5%）が認められたグループのペアに◇をつけた。以下の表でも同様である。

かなり高い肯定的回答を示したと見ることができ。自由記述においては高等専門学校からも、解釈や自己評価の難しい評価基準があったことは指摘されているが、全体的に妥当な基準であったという記述もなされている。前述のように、高等専門学校の認証評価は工学系のほぼ均質な機関を対象としており、分野や特徴に適合した基準にしやすい、また多くのプログラムは日本技術者教育認定機構（JABEE）の評価を受けていることで、対象校側にもPDCAのマネジメントサイクルへの理解がなされていると推察される。一方、法科大学院については、大学評価基準およびその解釈に関する批判が自由記述でいくつか寄せられている。

4.2 評価の方法および内容

問2は、評価の方法および内容の適切性について、(1) 自己評価、(2) 訪問調査、(3) 意見の

⁵ 等分散が認められた場合は最小有意差の検定を行い、等分散が認められなかった場合はTamhaneのT2検定を行った。

申立て、の3つの過程に分けて問うている。図2に示すように、全体として多くの項目で肯定的回答であり、4を超えるものも多い。全体的な状況としては、各対象校ともに自己評価を適切に実施できたと考え(2-1①)、訪問調査については事前の機構とのやりとり(2-2①, ②)も、当日の内容(2-2③, ④)も適切であったととらえられている。意見の申立ても実施方法としては適切と考えられている(2-3①, ②)。

一方、平均値が低い回答としては、自己評価に用いる資料が事前に十分に蓄積されていないこと(2-1②)や、資料の選択に困難を感じた対象校が多いこと(2-1③)が示されている。また、特に大学・短期大学および高等専門学校という機関単位の評価では自己評価書の文字数制限が制約となっ

たと考えている対象校が多くあった(2-1⑥)。

対象校間の比較(表4)では、まず、「自己評価書の完成度に満足しているか」(2-1⑤)について、大学・短期大学が最も高く、法科大学院に比して有意に差がある⁶。大学・短期大学のうちで多くを占める国立大学は既に試行的評価や法人年度評価において何度かの評価を経験していることや、一部の大学には評価対応専門の部署が存在しているという背景もあり⁷、比較的満足できる自己評価を行えたと思定される。その一方で、法科大学院は自己評価書の文字数制限は十分であると考え(2-1⑥)、組織単位が小さいために評価に必要な資料は機関レベルよりは明確であり(2-1②)、かつ、蓄積されていたようであるが(2-1③)、それでも自己評価書の満足の程度は相対的には低くなっている。自由記述では、法科大学院の評価が適格認定であることにより、自己評価書を教育研究の改善よりは基準への適合性を示すために作成した意識が強く、評価書の作成行為自体への満足度が低いことがうかがえる。

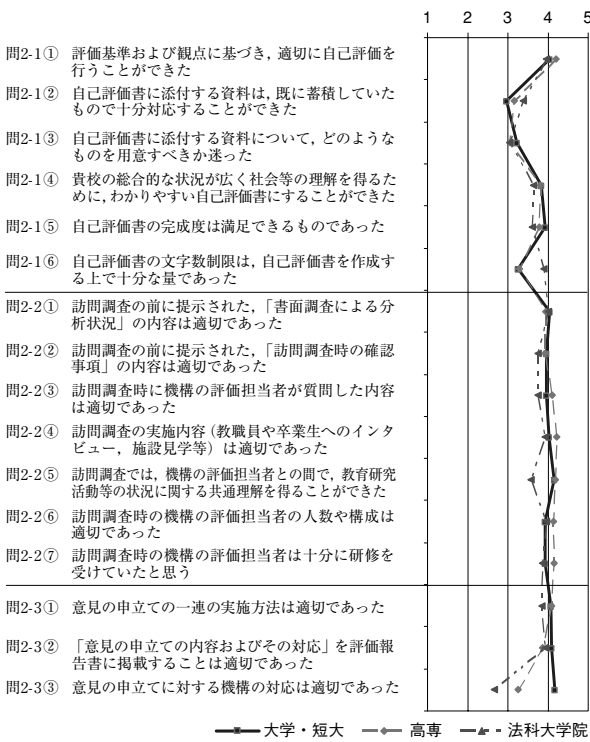


図2 評価方法および内容についての項目ごとの平均値

表4 評価の方法および内容についての分散分析の結果

項目	F 値	グループ1		グループ2		グループ3		多重比較		
		大学・短大		高専		法科大学院		グループのペア		
		平均	N	平均	N	平均	N	1と2	1と3	2と3
問2-1⑤ 自己評価書の完成度への満足	2.392 [†]	3.93	54	3.79	56	3.61	28		◇	
問2-1⑥ 自己評価書の文字数制限は適切	4.075*	3.26	54	3.29	56	3.89	28		◇	◇
問2-2③ 評価担当者の質問内容は適切	2.848 [†]	3.96	53	4.11	56	3.75	28			
問2-2⑤ 評価担当者共通理解を得た	7.808**	4.15	54	4.18	56	3.57	28		◇	◇

** p<.01, * p<.05, † p<.10

⁶ 本稿では分析対象としないが、評価者へのアンケートにおいても、大学・短大の評価者が自己評価書の質に関する質問に最も高く回答している。関連する質問に対する大学・短期大学(N=149)、高等専門学校(N=130)、法科大学院(N=85)の各評価者の回答平均値を以下に順に示す。「自己評価書は理解しやすかった」=3.62, 3.17, 3.32。「自己評価書には評価基準および観点の内容が適切に記述されていた」=3.66, 3.36, 3.26。「自己評価書には必要な根拠資料が引用・添付されていた」=3.54, 3.25, 3.26。評価者の回答の詳細については、大学評価・学位授与機構(2007a, b, 2009)を参照。

⁷ アンケートでは問8において対象校内部の評価実施体制を自由記述で問うており、大学の回答からは評価を専門に扱う常設部署や委員会組織が設置されている場合が多いことが示されている。

訪問調査については、「評価担当者と共通理解を得たか」(2-2⑤)について、高等専門学校および大学・短期大学では平均4以上という高い回答であり、訪問調査の有効性を認めている。訪問調査では、公式な面談を行った後に、評価者の私見を含む非公式な意見交換が行われることが常となっているが、それが有効であったとの自由記述もみられる。機構の認証評価では評価結果はあくまでも評価委員会によって保証された公式な見解となり公表されるが、それとは別に、この意見交換の機会があたかも外部評価のような役割を担っている場合もあると言える⁸。同質問に法科大学院は3.57と、肯定的ではあるが有意に他の学校種よりも低い回答であった。訪問調査が共通理解を得る場としては十分でなく、評価担当者との理解の相違が残った可能性がある⁹。

意見の申立てについては、申立てを行った対象校のみの回答を集計した。「意見の申立てへの対応が適切になされたか」(2-3③)について、大学・短期大学が、高等専門学校および法科大学院よりも高い。この項目に関しては、法科大学院は予備評価・本評価ともに「基準を満たさない」となった対象校があり、それに対する申立てが受け入れられなかった場合には、対立的な関係が生じている。これをどのように考えるかは見解の相違があるだろう。適切性が低いという回答を表面的にみれば評価結果への対象校からの不満があると見える。しかし、質保証という観点から見れば、既に設定された設置基準や評価基準に合致しなければ低い評価結果を生むことは当然と言える。その点で、評価基準の内容やその解釈について評価機関と対象校とで十分なコンセンサスを形成することが重要であり、さらなる検討を要する。

4.3 評価の作業量

評価の作業量については、(1)評価に費やした

作業量、(2)評価作業に費やした労力についての項目からなる¹⁰(図3)。

作業量については、自己評価書の作成(3-1①)の作業量が平均4.5で全体的に「大きい」という回答である¹¹。訪問調査や意見申立ても作業量はあるが、特筆すべき状況ではない。一方、評価作業に費やした労力が認証評価の3つの目的それぞれに見合うものであったかという質問(3-2①~③)に対しては、対象校グループで違いはあるが、ある程度の肯定的な回答が寄せられている。すなわち、作業量の負担感はあるが、可能な限りの軽減を今後行うことは必要であるが、評価目的実現を損なわないような適切なバランスも考慮することが求められる。

分散分析によるグループ間の比較では、訪問調査での確認事項や事前準備について、大学・短期大学の方が他グループに比べて対応の作業量は少なく感じており、大学・短期大学では訪問調査での確認事項を極力抑えていることが影響したと考えられる。また、自己評価のための作業量も大学は比較的に少なく回答しており、前述のように試

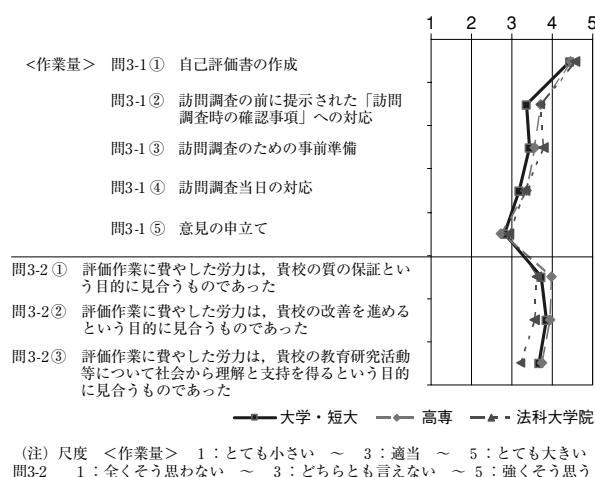


図3 評価の作業量についての項目ごとの平均値

⁸ これは認証評価に含まれる活動ではないため、機構による公式な検証では十分に触れられていない。

⁹ 評価者へのアンケートにおいても、法科大学院の評価者は訪問調査の効果について肯定的ではあるが相対的に低い回答である。関連する質問に対する大学・短期大学(N=144)、高等専門学校(N=130)、法科大学院(N=84)の各評価者の回答平均値を以下に順に示す。「訪問調査によって不明な点を十分に確認することができた」=4.30, 4.17, 4.01。「訪問調査では対象校と、教育研究活動等の状況に関する共通理解を得ることができた」=4.08, 4.06, 3.62。

¹⁰ アンケートでは実施時期の適切性についても問うているが2択質問のため省略する。

¹¹ 作業量の設問は3が「適切」という設定であり、たとえ「大きい」および「小さい」の両方の回答が多い双峰性の分布でも平均値は3に近くなる。ただし、本調査結果では双峰性は見られず、3以上の回答に偏っていた。

表5 評価の作業量についての分散分析の結果

項目	F 値	グループ1		グループ2		グループ3		多重比較		
		大学・短大		高専		法科大学院		グループのペア		
		平均	N	平均	N	平均	N	1と2	1と3	2と3
問3-1② (作業量) 訪問調査時の確認事項への対応	4.117*	3.37	54	3.71	56	3.71	28	◇	◇	
問3-1③ (作業量) 訪問調査の事前準備	3.143*	3.43	54	3.57	56	3.79	28		◇	
問3-2① 評価作業の労力は質保証という目的に見合う	2.530 [†]	3.72	53	3.98	56	3.61	28			◇
問3-2③ 社会からの理解・支持を得るといふ目的に見合う	4.083*	3.67	52	3.75	56	3.21	28		◇	◇

* p<.05, [†] p<.10

表6 説明会・研修会についての分散分析の結果

項目	F 値	グループ1		グループ2		グループ3		多重比較		
		大学・短大		高専		法科大学院		グループのペア		
		平均	N	平均	N	平均	N	1と2	1と3	2と3
問4③ 説明会の内容は役に立つ	2.844 [†]	4.06	54	4.23	56	3.85	27			◇
問4④ 理解しやすい配布資料	2.834 [†]	3.80	54	4.04	56	3.78	27	◇		
問4⑤ 研修会の内容への理解	2.577 [†]	3.81	54	4.07	55	3.89	27	◇		
問4⑥ 研修会の内容は役に立つ	6.301**	4.06	54	4.21	56	3.63	27			◇
問4⑧ 訪問説明は役に立った	8.303**	4.29	48	4.40	55	3.78	27		◇	◇
問4⑨ 機構の対応は適切	3.177*	4.26	54	4.36	56	4.00	27			◇

** p<.01, * p<.05, [†] p<.10

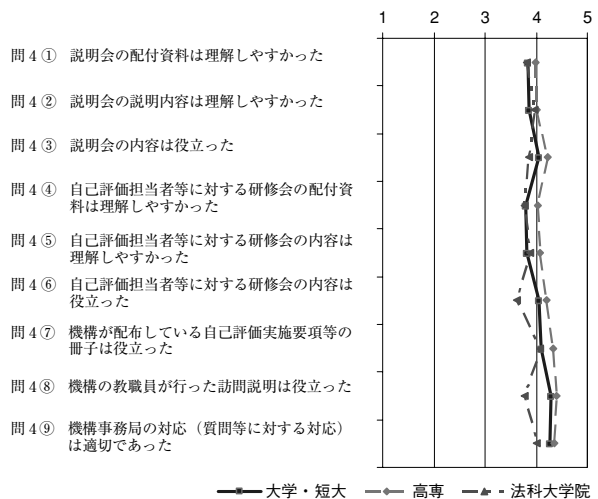
行的評価の経験や学内の評価対応体制の構築が関係していると考えられる。

評価作業に費やした労力の費用対効果の点については、全体的に、高等専門学校、大学・短期大学、法科大学院の順に高い回答であった。高等専門学校は規模的にも小さく、作業量自体の認識は最も少ないわけではないが、目的に対応する効果が得られたことから負担の適正感を感じているとみられる。法科大学院は評価対象の規模は最も小さいが、負担感が最も高いという結果となった。

4.4 評価に関する説明会・研修会

機構では、対象校に対して、評価の方法等に関する説明会や研修会を定期的で開催しており、さらに、多くの対象校を個別に訪問して説明を行っている。図4に示すとおり、これらに対する対象校からの回答は全ての項目において平均3.5を上回り、全体的に評価が高い。

分散分析の結果は表6であり、高等専門学校が全ての項目で4を超える極めて高い回答となっており、機構からの説明が特に適切に機能したと考えられる。大学・短期大学と比して高等専門学校では対象校がほぼ均質であるため、評価のための資料や研修会での説明が理解しやすい適切なものできたことが正の効果を及ぼしているものと推



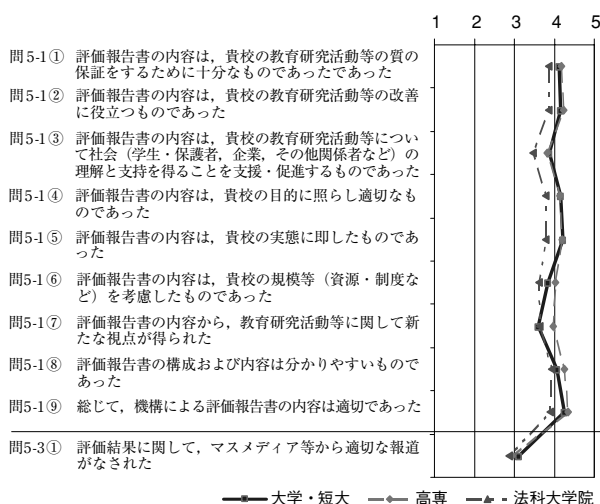
(注) 尺度 1: 全くそう思わない ~ 3: どちらとも言えない ~ 5: 強く思う

図4 説明会・研修会についての項目ごとの平均値

察される。

4.5 評価結果 (評価報告書) について

問5は、評価結果である評価報告書について問うている。質問は(1) 評価報告書の内容の妥当性、(2) 自己評価書および評価報告書の対象校での公表の有無、(3) 評価結果に関するマスメディア等の報道の適切性、のサブカテゴリから構成されている。このうちの(2)は2択質問である



(注) 尺度 1: 全くそう思わない ~ 3: どちらとも言えない ~ 5: 強くそう思う

図5 評価結果(評価報告書)についての項目ごとの平均値

ため分析から除外する。

(1) の回答結果は図5に示す通り、ほぼ全ての項目で3.5以上を上回り、対象校が評価結果におおむね納得していると言える。その中でも平均値が低い質問は、「評価報告書の内容が社会の理解と指示を促進するものか」(5-1③)であり、特に法科大学院では回答が3.5を下回った。機構が行った3年間の認証評価のうち「基準に適合しない」という評価結果が生じたのは法科大学院のみであり、自由記述では、特定の基準への不適合によって評価結果全体が不適合となることにより、社会からの誤解が生じる問題が指摘されている(村岡 2008も参照)。

また、(3) のマスメディア等からの報道の適切性についても、いずれの対象校グループとも3程度の低い値である。自由記述からは、いくつかのメディアが評価についての理解が不十分なまま記事を書いていると対象校は認識しており、評価結果が悪い箇所に過度に注目した報道がなされる傾向があることから、対象校に対する社会からの支援を逆に損なう可能性が指摘されている。特に3年目の評価結果については、基準に不適合という評価結果となった法科大学院に報道が終始し、機関別認証評価は報道がほとんどなされておらず、機構によるさらなる広報への要望も寄せられている。

グループ間の回答比較を表7に示す。評価報告書が対象校の目的や実態に即しているか(5-1④、

表7 評価結果(評価報告書)についての分散分析の結果

項目	F 値	グループ1		グループ2		グループ3		多重比較		
		大学・短大		高専		法科大学院		グループのペア		
		平均	N	平均	N	平均	N	1と2	1と3	2と3
問5-1② 改善に役立つ	3.336*	4.15	54	4.23	56	3.86	28			◇
問5-1④ 貴校の目的に適切	3.586*	4.15	54	4.14	56	3.79	28		◇	◇
問5-1⑤ 実態に即している	4.942**	4.22	54	4.20	56	3.79	28		◇	◇
問5-1⑥ 規模等の考慮	2.732†	3.85	54	4.02	56	3.61	28			◇
問5-1⑦ 新たな視点を得る	4.226*	3.59	54	3.98	56	3.64	28	◇		
問5-1⑧ 構成および内容は分かりやすい	3.255*	4.06	54	4.25	56	3.93	28			◇
問5-1⑨ 総じて、内容は適切	6.000**	4.26	54	4.34	56	3.89	28			◇

** p<.01, * p<.05, † p<.10

⑤), 改善に役立つか(5-1①)という質問に対して、法科大学院の回答は3.8程度と決して低い値ではないが、高等専門学校や大学・短期大学では4.2前後と非常に高い値を示している。前述のように、高等専門学校や大学では訪問調査において評価者と対象校の間で共通理解が形成されたという認識が高く、評価報告書の内容にも対象校は納得していると考えられる。

一方、「評価報告書から新たな視点が得られたか」(5-1⑦)については、大学・短期大学が最も低い値(3.59)となり、高等専門学校との間で有意差が見られた。大学・短期大学は評価報告書の内容は適切なものであったと考えていたが、その内容の多くは新たな視点よりは、自己評価の段階で自らも把握したものであったと推察される。前項のアンケート結果では、大学・短期大学では対象校自身が満足できる自己評価書の作成が行われていることが示されており、そのために第三者評価者から全く新たな視点の指摘がなされる可能性が下がっていることが考えられる。

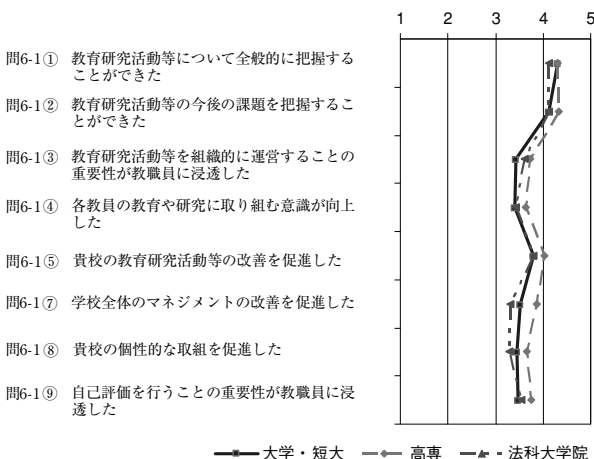
4.6 評価を受けたことによる効果・影響について

問6は評価の効果について、(1)対象校において自己評価を行ったことによる効果・影響、(2)

機構からの評価結果（評価報告書）を受けての効果・影響、にわけて問うている¹²。

図6で示すように、自己評価過程において「教育研究活動全般」（6-1①）やその「課題」（6-1②）を把握することができたという回答が、いずれの対象校グループも4を超えており、「教育研究活動の改善を促進」（6-1⑤）も3.8以上の高い回答である。すなわち、定められた評価基準・観点のもとで自己評価を行うことによって、教育研究活動の課題を把握し、改善に結びつけた対象校が多い。これ以外の「組織的運営の重要性の浸透」「教員の教育研究への意識向上」「個性的取組の促進」などの項目は3.5～3.6程度であり、ある程度の効果が得られたという結果にとどまる。高等専門学校の自由記述では「若い教職員が運営に参加する意欲がもてるようになる」などの効果の指摘はあるが、それでも、評価作業に携わっていない教職員には意識改革が生じた状態にはなっておらず、今後の課題として認識している対象校も多い。

一方、図7は（2）機構の評価結果を受けての効果・影響を示しており、（1）の自己評価による効果・影響と同様の質問を含む。結果は全体的には図6と同様の傾向である。図6と図7の結果を比べれば、上述の「教育研究活動全般」や「課題」の把握については、自己評価過程のほうがいずれの対象校グループの回答平均値も高い。すなわち、

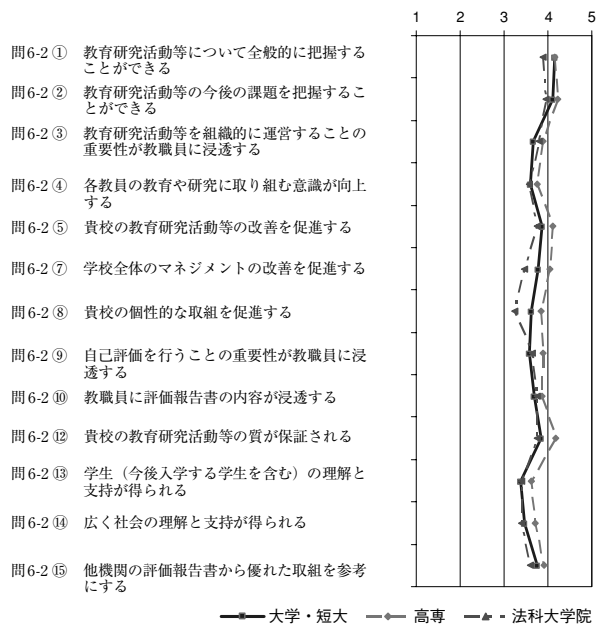


(注) 尺度 1: 全くそう思わない ~ 3: どちらとも言えない ~ 5: 強くそう思う

図6 自己評価による効果・影響についての項目ごとの平均値

実態の把握には自己評価過程が重要であり、外部の評価者が短期間で把握し指摘できる内容は限られる。一方、その他の項目については、法科大学院での「教育研究活動の改善を促進」「個性的取組の促進」をのぞいた全てにおいて、機構の評価報告書を受けた効果の方が高い結果となった。いずれの対象校グループで2つの回答差が統計的に有意である質問項目を表8に示す。「教職員の意識向上」（6-2③）は全ての対象校グループにおいて5%有意で機構の評価報告書を受けた効果のほうが高く、「組織的運営の重要性の浸透」「マネジメントの改善促進」「個性的取り組み促進」は大学・短期大学と高等専門学校において同様に5%水準で有意に高い。組織の運営改善に向けて教職員の意識変化を引き起こすには、学内での取り組みに加えて、外部から指摘されることが一つの圧力やインセンティブとして機能しうることを示していると考えられる。

対象校グループ間の比較では、いずれの項目においても、他グループに比べ高等専門学校の回答が高い。分散分析の結果は表9に示しており、「個



(注) 尺度 1: 全くそう思わない ~ 3: どちらとも言えない ~ 5: 強くそう思う

図7 評価結果による効果・影響についての項目ごとの平均値

¹² アンケートでは具体的な改善内容を記述していただいている。その内容については、各年度の報告書（大学評価・学位授与機構 2007a, 2007b, 2009）を参考。

表8 自己評価と評価結果での効果・影響の差異のt検定

項目	大学・短大		高専		法科大学院	
	平均値の差	t値	平均値の差	t値	平均値の差	t値
① 教育研究活動等の全般的把握	-0.15	-2.06*	-0.14	-1.93 [†]	-0.21	-2.27*
③ 教育研究活動等を組織的に運営することの重要性が教職員に浸透	0.24	2.75**	0.16	2.43*	0.18	1.72 [†]
④ 各教員の教育や研究に取り組む意識向上	0.20	2.84**	0.13	2.18*	0.14	2.12*
⑦ 学校全体のマネジメントの改善促進	0.26	3.24**	0.20	3.67**	0.14	1.69
⑧ 個性的な取組促進	0.19	2.63*	0.18	2.46*	-0.04	-0.24
⑨ 自己評価を行うことの重要性の浸透	0.09	1.00	0.14	2.21*	0.11	1.00

** p<.01, * p<.05, [†] p<.10

視され、それに適合しない各校の工夫点等が尊重されない結果になった旨が主張されている。一方、高等専門学校は「マネジメントの改善促進」「教育研究活動等の改善促進」などの複数の項目で、法科大学院や大学・短期大学よりも高い。高等専門学校の自由記述に他グループと異なる特別な指摘は見られないが、評価を通じて長所や短所を把握し、改善につなげている旨が記されている。

4.7 まとめ

以上、各カテゴリーにおける大学・短期大学、高等専門学校、法科大学院の対象ごとの平均値の比較検定を行った結果について述べた。各対象校ごとの回答からは以下のような特徴がまとめられる。

大学はもっとも評価対象の規模が大きい組織であるが、試行的評価からの経験や、機関の評価専門部署の設置など実施体制の構築により、作業の負担軽減努力がなされ、作業量の認識は3グループ中で最も低い。また、自己評価書の満足度は最も高く、評価方法の適切性の認識や改善効果についても、ある程度の高い値は得られている。しかし、評価結果から新たな視点が得られる程度は対象校グループ中で最も低いなど、第三者評価者による評価に基づく新たな改善効果は相対的に見れば限定されたものになっている。

高等専門学校は多くの質問項目で最も高い回答を示している。高等専門学校にとっては、JABEEによる工学教育プログラム評価の経験はあるものの、機関レベルでの評価は新たな取り組みである。自由記述では、機構の説明会などを通じて「学生にとってよりよい高等専門学校作りのために評価を利用できることが認識されるようになった」旨や、評価において「取り組みの良いところを見ようという姿勢がうかがえた」ことを肯定する内容もあり、それぞれの高等専門学校の長所を踏まえて改善を促進しようとする意識が高等専門学校と評価者とで共有されたことがうかがえる。そのため、評価実施による効果の認識も高かったと考えられる。

一方、法科大学院は、評価対象の組織の規模は最も小さいが、機関単位とは異なり、組織内部で評価実施に十分な人員や資金を措置しにくいために、評価における作業量は3グループ中で最も多

表9 自己評価・評価結果による効果・影響についての分散分析の結果

項目	F値	グループ1		グループ2		グループ3		多重比較		
		大学・短大		高専		法科大学院		グループのペア		
		平均	N	平均	N	平均	N	1と2	1と3	2と3
問6-1⑦ マネジメントの改善促進	6.701**	3.52	54	3.86	56	3.32	28	◇		◇
問6-1⑧ 個性的な取組の促進	2.733 [†]	3.44	54	3.66	56	3.29	28			◇
問6-2⑤ 教育研究活動等の改善促進	3.377*	3.87	54	4.11	56	3.75	28			◇
問6-2⑦ マネジメントの改善促進	7.385**	3.78	54	4.05	56	3.46	28	◇		◇
問6-2⑧ 個性的な取組の促進	5.011**	3.63	54	3.84	56	3.25	28			◇
問6-2⑨ 重要性の浸透	2.491 [†]	3.57	54	3.89	56	3.64	28	◇		◇
問6-2⑫ 質が保証される	4.623*	3.85	54	4.18	56	3.75	28	◇		◇

** p<.01, * p<.05, [†] p<.10

性的な取り組みの促進」については(1),(2)ともに法科大学院は3.5を下回る回答となり、高等専門学校との間で有意差がある。法科大学院からの自由記述においては、評価基準への整合性が重

く認識されている。評価作業に費やした労力が質保証や改善促進といった目的に見合うものであったという質問に対しても、他グループに比べ回答平均値が低い。また、適格認定という性質から評価基準への適合度が重視されるため、評価者と大学とで解釈の相違がある場合に、それが評価結果に大きく影響することになり、評価結果の妥当性の認識なども比較的に低めになっている。特に個人的な取り組み促進という効果への回答が低くなり、適格性の保証と個性促進との両立が容易に行い得るものではないことを示していると言えよう。

5. 構造分析

以上のように個々の質問項目ごとの回答が得られたが、では、それぞれの質問がいかなる関係にあり、最終的に対象校の改善という効果が実現されるのか。また、対象校の違いによってその過程に差異はあるのか。これらの問いを検討するために、共分散構造分析を用いて改善促進へと結びつく因果構造のモデルを検証する。はじめに、観測変数によって規定される潜在変数を作成するために、質問項目の探索的因子分析を行い、因子構造を明らかにする。その結果に基づき、評価結果による改善促進モデルを構築して共分散構造分析を行う。

5.1 探索的因子分析

まず、質問項目のカテゴリーごとに探索的因子分析（主因子法、斜交プロマックス回転）を行う。ただし、「4.3 評価の作業量」は、回答の欠損値が多いため以下の分析から除外する。因子の抽出に際しては、固有値1以上を基準に行い、その結果をほとんどの場合には解釈可能であった。ただし、問6の「機構の評価結果を受けたことによる効果・影響」については因子数が3となったが、各因子の解釈が困難であったため、因子数を4に指定した。探索的因子分析の結果は巻末の別表A-1～A-6に示す。結果、それぞれの質問項目カテゴリーにおいて表10に示すような因子を抽出し、それぞれに名称をつけた。

5.2 因果モデルの検証

探索的因子分析の結果から得られた因子を用いて、因果構造のモデルを構築する。アンケートの

表10 探索的因子分析により得られた因子

質問項目のカテゴリー	得られた因子の名称	共分散構造分析におけるカテゴリー
問1 評価基準および観点	因子1：評価基準および観 点の適切さ	インプット／プロ セス
	因子1：訪問調査の適切さ	
問2 評価の方法および内容	因子2：自己評価書の満足	アウトプット
	因子3：訪問調査前の提示 内容の適切さ	インプット／プロ セス
問4 説明会 ・研修会	因子1：説明・研修会の理 解	インプット／プロ セス
	因子2：機構の有効な対応	
問5 評価結果 (評価報告書)	因子1：評価報告書の満足	アウトプット
問6-1 自己評 価による効果 ・影響	因子1：意識変化	アウトカム
	因子2：現状把握および改 善促進	
問6-2 機構の評 価結果を受けた ことによる効果 ・影響	因子1：意識変化	アウトカム
	因子2：現状把握	
	因子3：理解と支持	
	因子4：改善促進	

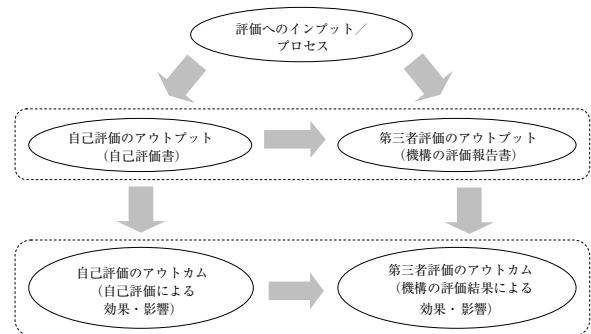


図8 評価のメカニズム解明のための理論モデル

質問項目は、評価実施の目的、方法、結果、効果というように、評価作業の時間的な流れに即している。そのため、図8に示すように、評価のインプットやプロセスである評価基準や評価方法の適切性が、大学が行う自己評価のアウトプットである自己評価書、ならびに機構が行う第三者評価のアウトプットである評価報告書の妥当性に影響すると仮定する。さらに、これらのアウトプットの適切性が、アウトカムである評価による効果・影響の程度に関係すると仮定する。また、評価は、大学での自己評価作業と機構での評価者による第三者評価作業という、大きく2つのプロセスがあり、前者が時間的に先行して後者の作業の基礎となる。そのため自己評価のアウトプットとアウトカムは、それぞれ、第三者評価のアウトプットと

アウトカムにも影響するという関連性を仮定した。

このような概略的なモデルを基礎にして、より詳細に、インプット、プロセス、アウトプット、アウトカムの要素がどのような関係にあるかを共分散構造分析を用いて検証する。各質問項目がインプット等のどの概念の要素と考えられるかは、前述の表10に示した。ただし、問6-1の自己評価のアウトカムは因子分析の結果では因子数が2つであったが、問6-2の第三者評価のアウトカムとの対応関係を考慮して、第二因子を問6-2と同様となるように「自己評価による現状把握」と「自己評価による改善促進」とにさらに区分して構成概念として用いた。概念変数における観測変数に関しては、因子負荷量が0.55以上の上位最大3つを用いた。その際、因子負荷量が0.7を超える高い値の項目が複数ある場合には、構成概念を説明しやすい観測変数を選択した。なお、分析に用いる標本数は、欠損値のあるケースを除外した結果、計121校（大学・短期大学44校、高等専門学校52校、法科大学院25校）となった。

これらの変数を用いて、評価作業の時間的な流れに沿って因果関係を仮定し、有意でないパスを削除していった結果、図9に示す構造方程式初期モデルを得た。この初期モデルの適合度指標は、 $GFI = 0.794$, $AGFI = 0.751$, $RMSEA = 0.052$ である。通常、 GFI ならびに $AGFI$ は1に近いほど説明率が高く、 $RMSEA$ が0.05以下であれば適合度が高く、0.10以上であればモデルを採用しない（豊田1998）。本分析の結果は、受容するに微妙な適合度である。また、狩野（1997）はモデルの適合度判断において数百程度の標本数であればカイ2乗検定、1000前後以上であれば GFI など他の指標を適合性の検定に用いるのが妥当であるとしている。本研究では欠損値を除いた分析データ数が121であったので、 GFI などの適合度を参考にしつつ、最終的にカイ2乗検定によってモデルの採択を決めることが望まれる。図9のカイ2乗検定による確率は0.00であり、有意水準0.05でこの初期モデルは棄却されることになる。

そのため、このモデルを何らかの形で修正する必要がある。初期モデルはインプットからアウトカムまでを一つの包括的なモデルとして組み立てており、観測変数の数は30を超える。そのため多数の変数間の関係の有無を仮定していることにな

り、モデル全体が複雑で、適合度は低くなる傾向にある。豊田（1998）は必然性のある場合を除いては観測変数の数を30以内に収めるように指摘している。そのため、図9の初期モデルを以下のように上下の二つに分割して検討する。

一つはインプット／プロセスからアウトプットまでのモデルである。前述の図9の破線の上位に位置する部分のみ（最上部から「自己評価書の満足」、「機構の評価報告書の満足」まで）をそのまま取り出し、分析を行うと、ほぼ同様のパス係数が得られ、 $GFI = 0.897$, $AGFI = 0.857$, $RMSEA = 0.031$ であり、カイ2乗検定による確率は0.185と受容できるモデルとなった。このモデルによれば、アウトプットの一つである自己評価書の満足度には、評価基準および観点が適切であると対象校が考えているか、および、機構が訪問説明などを適切に行ったかが影響している。また、機構の評価報告書の満足度には、自己評価書の満足度に加えて、訪問調査が適切に行われたと考えているかが影響する。

もう一つのモデルは、図9の破線の下に位置する、アウトプットからアウトカムまでを基本とするものである。この中で、「学生・社会からの理解支持」を削除して、評価による改善効果に焦点をあてたモデルとして検証を行った。結果は、パス係数は図9と大きくは変化せず、適合度は $GFI = 0.898$, $AGFI = 0.850$, $RMSEA = 0.043$ と良好な値になり、さらに、カイ2乗検定による確率が0.061と有意水準0.05で受容できるモデルとなった。パス係数は「機構の評価結果による現状把握」から「機構の評価結果による意識変化」のみ有意でないが、それ以外はすべて統計的に有意である。

この二つ目のモデルの結果によれば、自己評価により現状把握がなされるか否かは、満足できる自己評価書が作成されたか否かによって強く規定される。さらに、現状把握が行われたことで、教職員における組織的運営の重要性や自己評価実施の重要性に関する意識が変化し、これら現状把握と意識変化があいまって、教育研究活動等の改善促進へと強くつながっていく。

一方、機構の評価報告書から現状把握がなされる程度は、自己評価段階で自ら現状把握をした程度と、評価報告書が満足できるものであったかによって、やや弱く規定されている。しかし、評価

報告書を受け取った後の教職員の意識変化は、既に自己評価段階において意識変化がなされた程度によって強く規定されており、機構の評価報告書から現状把握がなされたかには統計的に有意には影響されない。また、評価報告書を受け取った

後の改善促進は、評価報告書に基づく意識変化、自己評価過程での改善促進、評価報告書により現状把握がなされた程度の3つから総合的に影響を受けている。

この結果からは、対象校自身が満足できるよう

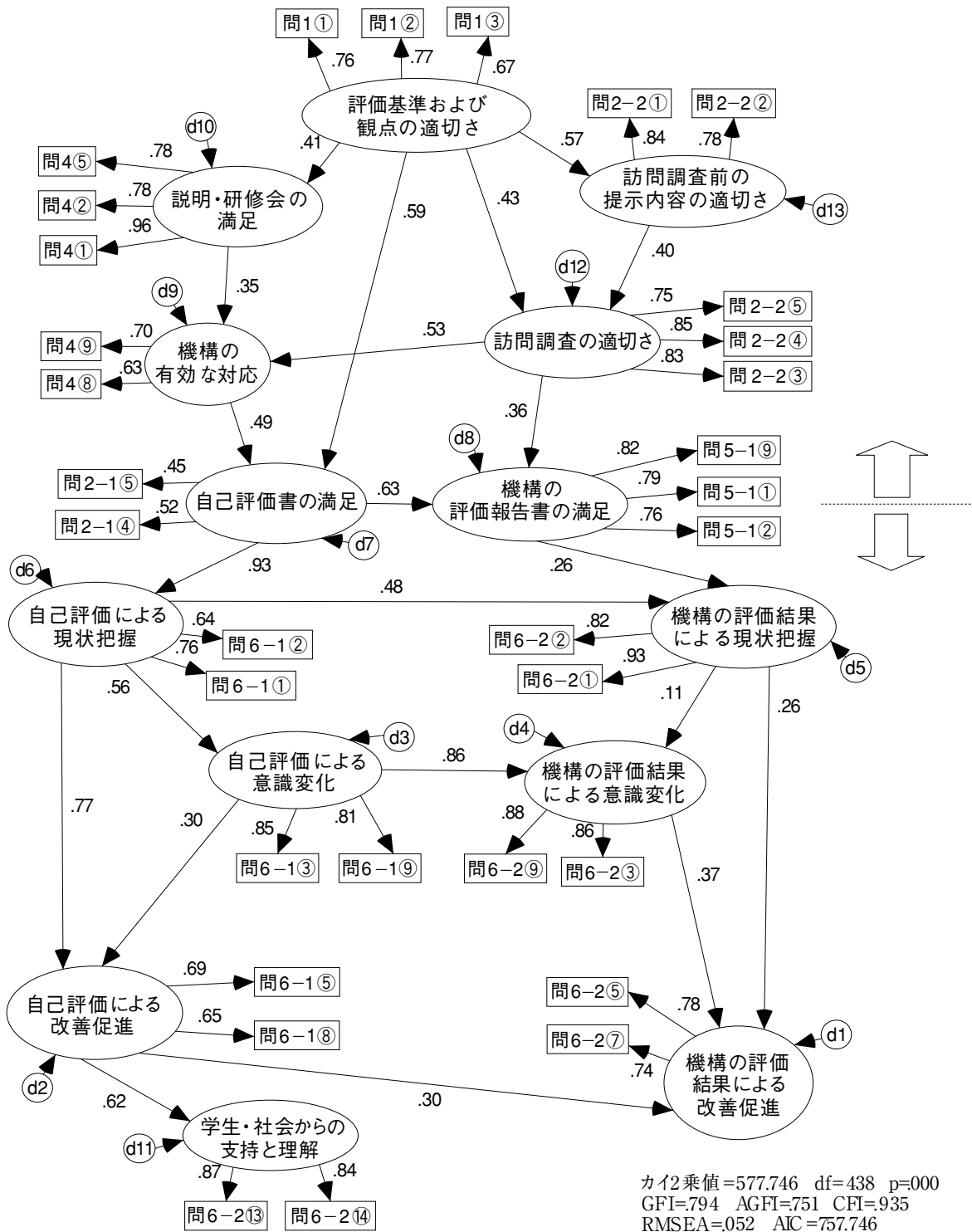


図9 評価結果による改善促進の初期モデル (標準化解)

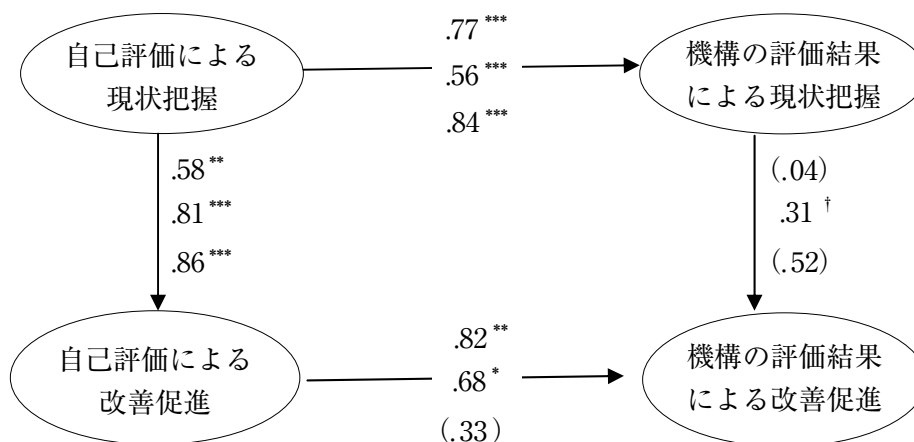
な自己評価を行い、それによって現状把握を適切に行うことによって、改善促進へと結びつくというパスが存在することを示しており、評価の効果は対象校自身の取り組みによって大きく影響されると言えよう。一方、機構の第三者評価は、外部の評価者が知り得る情報もそのための時間も大学自身と比べれば限定されたものとならざるを得ないため、自己評価を大きく超える現状把握が新たに行われることは難しい。そのため、評価結果を受け取った後の改善は、評価報告書による現状把握だけでなく、自己評価段階で既に改善がなされる程度、および、評価結果が一つの外圧やインセンティブとして教職員の意識変化を促している程度の3つが総合的に影響を与えることを示している。

5.3 対象校間の比較 (多母集団同時分析による因果関係の比較)

次に、大学・短期大学、高等専門学校、法科大学院の3グループで改善に至る構造に違いがあるかを検討するために多母集団分析を行う。多母集団同時分析の際には、同一の因子構造をもつモデルをグループごとに個別分析し、適合度からモデルが採用されることを前提としている。しかし、3つのグループに分けることによりサンプル数が減り、図9の下部モデルを用いた個別分析は適合度が悪くなるため、さらに分析の焦点を絞ったモデルで検討を行う。すなわち、自己評価と第三者

評価のアウトカム部分に焦点をおき、さらに「自己評価による意識変化」と「評価結果による意識変化」が不適解となるために削除して単純なモデルとし、3グループにおける適合度に問題がないことを確認したうえで、多母集団同時分析を行った。図10は多母集団同時分析により推定された因果関係を示しており、係数は上から大学・短期大学、高等専門学校、法科大学院の順に示している。モデルの適合度は $GFI = 0.905$, $AGFI = 0.771$, $RMSEA = 0.035$ であり、またカイ2乗検定による確率は0.235であり、採用できるモデルである。

3つのグループにおけるパス係数の差の検定の結果では、構成概念間のどのパスにおいても有意差は認められなかったが、それぞれの標準化解による特徴は現れた。まず、大学・短期大学においては、「自己評価による改善促進」が「機構の評価結果による改善促進」にもっとも強い影響を与えており (因果係数0.82)、遡ると「自己評価による現状把握」が「自己評価による改善促進」を強く規定する (因果係数0.58)。一方、「自己評価による現状把握」は「機構の評価結果による現状把握」へ強く影響はするが、そこから「機構の評価結果による改善促進」への影響は統計的に有意でなかった。この結果から、大学・短期大学における認証評価は自己評価過程での教育研究活動等についての全般的な把握や今後の課題の把握が重要であり、それにより、教育研究活動等の促進や個性的な取り組みが促進されると解釈できる。



注) 上段: 大学・短期大学, 中段: 高等専門学校, 下段: 法科大学院

*** $p < .001$ ** $p < .01$ * $p < .05$ † $p < .10$, () 付きの係数は有意差なし

図10 アウトカム内の評価結果による改善促進モデルの多母集団同時分析 (標準化解)

高等専門学校においては、大学・短期大学と同様に「自己評価による現状把握」から「自己評価による改善促進」への直接的な影響力が大きく、そこから「機構の評価結果による改善促進」への影響力を有している。さらに、「自己評価による現状把握」から「機構の評価結果による現状把握」へ影響を及ぼし、そこから「機構の評価結果による改善促進」へもやや影響を及ぼしている。つまり、自己評価過程だけでは改善効果は十分に生まれず、第三者評価による報告書の意義が大学・短期大学よりも大きい。

法科大学院においては、「自己評価による現状把握」から「自己評価による改善促進」と「機構の評価結果による現状把握」への直接的な影響はほぼ同等（因果係数はそれぞれ0.86と0.84）である。しかし、大学・短期大学と高等専門学校とは異なり「自己評価による改善促進」から「機構の評価結果による改善促進」への影響は統計的に有意でなく、また「機構の評価結果による現状把握」から「機構の評価結果による改善促進」も有意ではない。すなわち、最終的な改善効果は別の要因が影響していると考えられ、特に評価結果が「基準に適合」か「不適合」かによって、評価報告書の適切性やその効果についてのアンケート回答も大きく異なっており、評価結果自体に影響される部分があると考えられる。

6. 議論

以上のように、本稿では、大学・短期大学、高

等専門学校、法科大学院という3つの認証評価の検証アンケート結果の比較を行うとともに、改善へ結びつく構造とその対象間の差異について検討した。これらの結果からは、評価対象の2つの属性—すなわち、1) 評価対象校での評価の内部浸透の程度、2) 評価対象のレベル（大きさ）—が、第三者評価の適切性や効果に大きな影響をもたらすと考えられる。以下、この2点について対象校ごとの特徴をまとめ、今後の機関評価および分野別評価の構築・修正のあり方への含意を検討してみたい。

一つ目の属性である「評価の内部浸透の程度」とは、対象校における評価経験の蓄積や自己評価体制の確立を基礎に、評価を自らの質向上のための作業として実施しうる能力の程度である。共分散構造分析の結果からは、最終的な改善促進は、自己評価過程やそれに基づく学内での改善活動に強く規定されていることが示された。すなわち、評価対象校が質の高い自己評価を行い、その結果を活用できる体制を有していることが、評価を有効なものとするのに重要である。ただし、アンケート回答の比較分析からは対象校の種別により評価の内部浸透の程度には差異があり、それによって評価システムの適切性の認識も異なっている。

概念的に整理するために、表11のような内部浸透の程度の段階区分を仮説として考えてみる。第1段階は、組織内部で自己評価や改善を組織的に行う体制が実質的に存在しない段階であり、評価

表11 評価の内部浸透段階仮説

	第1段階	第2段階	第3段階
対象校での評価の内部化の程度	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。	内部で常設の評価体制が設置される。	評価が日常的な改善体制と一体化する。
第三者評価の位置づけ	第三者評価を受けなければならないことが、教育・研究について組織的に考える初めての機会。	第三者評価は、内部での自己評価や教育改善を駆動させるための一つの外圧。	日常的に教育・研究活動の確認や修正が行われており、第三者評価への対応は、その総括や対外報告。
第三者評価システムの在り方	評価基準として教育・研究の運営で考慮すべき項目が詳細に示されていることが必要。自己評価作業が不慣れで十分ではないため、第三者としての評価者が長所・短所を詳細に指摘する必要がある。	評価や改善を進める一つの指針として評価基準が詳細に定められている必要はあるが、全ての内容を根拠をもとに詳細に確認して第三者評価を行う必要はない。内部質保証体制が機能していることを評価することにより、その構築を促進することに重点が置かれる。	全体的な状況については、大学内部の質保証機能の評価によって質保証を行う。大学自身が詳細な評価を行いたい高度な課題を設定し、専門知識を有する外部評価者によるコンサルテーション的な第三者評価を実施する必要。

を制度として強要されている段階である。次の第2段階は、組織内部に常設の評価体制が設置され、自己の改善のために評価を意図的に活用しようとする段階である。最後の第3段階は、評価が日常的な改善体制と一体化して特別な存在ではなくなる段階である。

本分析での3種類の対象校グループのうち、大学・短期大学は実際には国立大学が多く、自己点検・評価、「試行的」評価、国立大学の法人化による年度評価など、多種の評価の経験が蓄積されており、評価業務のための常設部署（いわゆるIR（インスティテューショナル・リサーチ）部門）を設置しているところも少なくない。そのため内部浸透の程度の第2段階に到達している大学もあると考えられる。大学・短期大学のアンケート回答の特徴は第2段階の特性と論理的に整合し、他の対象校グループと比して評価作業を効率的に行い、自己評価も自らが満足しうる質のものとなっていた。その一方で、改善へ機能しうる能力が対象校の内部に保有されるようになることで、第三者評価によって外部者から指摘される新たな視点は相対的に少なくなっていた。

一方、高等専門学校については、JABEEによる、5年間の準学士課程のうちの4年次以上の2年間と専攻科をあわせて一つの変則的な学士プログラムと見なした認定評価を受けている場合が多い。そのため、教育目標と教育内容の整合性の確保、PDCAサイクルの確立などには大学以上に慣れている場合も多いと考えられる。しかし、機関別認証評価の結果では、準学士課程と専攻科課程ごとに分けた目標設定の必要、一般科目の教育体制の改善、教育活動に対する学内評価制度の必要、AO（アドミッションズ・オフィス）入試への要請など、JABEEによるプログラム評価では十分に視野に入らない部分への改善要請が見られた。そのため、機関全体という点では、評価の内部浸透はいまだ第1段階の対象校も多く、多母集団分析の結果では、機構の評価結果に基づく現状把握によって改善が促進されることが示されていた。

法科大学院に関しては、組織自体が設立したばかりであるため、内部の評価体制の構築も今後の課題であり、第1段階にあると言える。そのため、現在は評価への負担感も高く、評価基準への理解やその内容への合意も対象校により差異がある。

今後は自ら定期的に点検と改善を図っていくことが望まれる段階にある。

このような段階仮説と各対象校の回答・分析結果からは、それぞれの段階において求められる評価システムは異なる可能性が示唆される。第1段階の状況にある対象校が多い場合の評価システムでは、その定義上、自己評価を行う際に着目すべき観点を対象校自らが設定することは困難である。そのため、第三者評価において詳細な評価基準を設定し、対象校との間で共通理解を形成することで、自己評価の質向上を支援することが重要となる。共分散構造分析の結果においても、機構が定めた評価基準が適切であり、訪問説明などの十分な対応・指導が行われていることによって、満足できる自己評価が行えることが示されていた。同時に、第1段階では対象校自身が各観点について自己分析を行う能力も十分ではないために、第三者評価者が詳細な評価を行って対象校自身が気づいていない長所・短所を指摘することが改善にとって重要になる。

しかし、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減減せざるをえない。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。

変更の方向性の検討は本分析の範疇を超えるものであるが、例えば考えられる一つの方向性としては英国や豪州のオーディット形式のように、対象校の内部質保証体制の評価に重点を移行させることが考えられる。現在の認証評価においても評価基準の一つとして「質の向上および改善のためのシステム」は設定されているが、さらに、評価者が教育内容や教育体制について一つ一つの観点を詳細に評価することはせず、対象校自身がそれを確認し修正する体制があることを二次的に確認することにいっそうの重点を置く方向である。また、もう一つ考えられる方向は、最低限の水準を超えるような、政策的課題への対応や国際的な水準への到達など、高度な内容を積極的に評価基準に含み、特定課題への専門的な見地からの助言を行うことである。これにより大学内部では得られない新たな視点が増すことになる。このように、

評価の内部浸透段階により適切な評価システムのあり方も変化すると考えられる。

評価システムの適切性や効果に影響する属性の二つ目には、評価対象のレベルが挙げられる。大学・短期大学および高等専門学校の認証評価は基本的には機関単位であり、法科大学院の評価は分野別のプログラム単位である。ただし、高等専門学校は多くが工学という単一分野であることや小規模であることから、プログラム単位の評価にも近い傾向を有していた。機関単位の評価では管理運営体制やFDなどの全学の質保証体制が重視され、対象校の規模や特性に十分に配慮するよう評価が実施されることが通常であり、また内部質保証構築を通じた自律性が尊重されやすい。そのため、対象校との間で大きな対立が生じることは少ないと考えられる。一方、プログラム単位の評価では科目構成や教員構成などの詳細な内容が評価対象となり、当該分野の教育・研究の内容に入り込んだ実質的な評価が可能となる。仮に対象校と評価者との間で対象校の現状や望まれる方向について共通理解が形成され、評価結果が納得して受容されれば、教育・研究の直接的な改善へとつながることになる。その反面、評価者と対象校とで教育・研究の内容や方法のあり方についての意見の相違も顕在化しやすく、対立構造が生まれやすい危険もある。実際、法科大学院のアンケート結果からは、評価者との意見の相違が訪問調査でも十分には埋まらなかった状況がうかがえる。法科大学院の認証評価は「適格認定」であることがさらに影響し、評価基準への適合の可否という評価結果が持つ意味が大きいために、評価者との意見の相違が論争を生じやすくしている。

このことから、分野別評価の設計は機関別評価以上に難しいものとなる可能性が示唆される。現在、中央教育審議会からの委託を受けて、日本学術会議において大学教育の分野別質保証の在り方についての審議が進められている。これらの審議の結果、仮に分野ごとに標準的な教育内容・方法が詳細に規定され、評価の運用時点で参照情報という位置づけ以上に厳格に適用されることがあれば、評価者と被評価側との対立構造が生まれやすくなると予想される。そのため、標準的な内容・方法をどこまで詳細に設定し、一方で大学の創意工夫をいかに適切な形で評価して推奨するかにつ

いて、十分な注意を払って分野別評価が設計・実施されることが望まれよう。また、適格認定のように評価基準への適合性が評価結果に大きく影響するような評価においては、評価以前から評価基準の解釈や大学の実態について情報交流を行うことにより、評価段階になって初めて合否が問われる状態はさけるべきであろう。この点においては、法科大学院の評価では本評価の前に予備評価が行われており、アンケートの自由記述では予備評価が本評価への参考になったという意見が複数見られる一方、予備評価と本評価とで判断が不統一であった旨の指摘もなされており、今後も本評価とそれ以前の評価や情報交換との関係をよりいっそう明確にしていく必要がある。

このように本分析では、機構がこれまで実施した異なる種類の認証評価の比較を行った。3種類と限られた数であることや、それらの間で評価対象や評価システムの複数の要因に差違があることから断定的な結論を行うことは難しいが、それでも以上のように評価対象の特性が評価システムの有効性に影響する可能性と今後への含意が示された。大学等のように変革を繰り返す組織を対象とする評価においては、一つの評価システムが、どのような対象にも何時でも適切であるということはありません。評価システムを常に変化させていかなければ、非効率なものに陥る可能性がある。認証評価制度では、評価基準や評価方法の変更は中央教育審議会の承認を得なければならないことから即時の修正には制約があるが、その中でも様々な対象の評価の適切性や効果を検証し、不断の変更を進めていくことが今後も求められよう。

参考文献

- 大学評価・学位授与機構 (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—』
http://www.niad.ac.jp/n_hyouka/jouhou/kenshou/1174681_959.html (2009/02/09アクセス)
- 林隆之, 齊藤貴浩 (2005) 「大学評価のメタ評価—評価は大学の研究活動を改善したのか?」『研究・技術計画学会 第20回年次学術大会講演要旨集』 pp.248-251.
- 齊藤貴浩, 林隆之 (2007) 「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』『日本評価研究』 vol.7 No.1, pp.33-46.
- 大学評価・学位授与機構 (2007a), 『平成17年度に実施した認証評価に関する検証結果報告書』
http://www.niad.ac.jp/n_hyouka/jouhou/1179217_989.html (2009/02/09アクセス)
- 大学評価・学位授与機構 (2007b), 『平成18年度に実施した認証評価に関する検証結果報告書』
http://www.niad.ac.jp/n_hyouka/jouhou/1179704_989.html (2009/02/09アクセス)
- 大学評価・学位授与機構 (2009), 『平成19年度に実施した認証評価に関する検証結果報告書』
http://www.niad.ac.jp/n_hyouka/jouhou/1180841_989.html (2009/02/09アクセス)
- 中央教育審議会 (2002), 『大学の質の保証に係る新たなシステムの構築について』
http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo0/toushin/020801.htm
(2009/02/09アクセス)
- 中央教育審議会 (2005), 『新時代の大学院教育—国際的に魅力ある大学院教育の構築に向けて—』
http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo0/toushin/05090501.htm
(2009/02/09アクセス)
- 中央教育審議会 (2005), 『学士課程教育の構築に向けて』
http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo0/toushin/1217067.htm
(2009/02/09アクセス)
- 羽田貴史 (2006) 「大学改革における評価の機能と役割」『京都大学高等教育研究』12号, 117-128.
- 村岡啓一 (2008) 「評価を受けて 一橋大学法科大学院」『IDE 現代の高等教育』2008年10月号, pp.11-14.
- 「特集 5年目を迎えた認証評価」『カレッジマネジメント』150号, 2008年5-6月.
- 豊田秀樹 (1998), 『共分散構造分析 (入門編) —構造方程式モデリング』, 朝倉書店, 東京.
- 狩野裕 (1997), 『AMOS EQS LISREL によるグラフィカル多変量解析—目で見る共分散構造分析—』, 現代数学社.

(受稿日 平成21年1月16日)

(受理日 平成21年2月5日)

別表 A-1 評価基準および観点に対する因子分析結果 (N=129)

質問項目	評価基準および 観点の適切さ
問1① 評価基準および観点の構成や内容は、貴校の教育研究活動等の質を保証するために適切であった	0.795
問1② 評価基準および観点の構成や内容は、貴校の教育研究活動等の改善を進めるために適切であった	0.741
問1③ 評価基準および観点の構成や内容は、貴校の教育研究活動等について社会（学生・保護者、企業、その他関係者など）から理解と支持を得るために適切であった	0.614
問1④ 評価基準および観点の構成や内容を、教育活動を中心に設定していることは適切であった	0.366

別表 A-2 評価の方法および内容に対する因子分析結果 (N=129)

質問項目	第1因子	第2因子	第3因子
	訪問調査の 適切さ	自己評価書 の満足	訪問調査前 の提示内容 の適切さ
問2-2④ 訪問調査の実施内容（教職員や卒業生へのインタビュー、施設見学等）は適切であった	0.954	-0.060	-0.095
問2-2③ 訪問調査時に機構の評価担当者が質問した内容は適切であった	0.772	-0.075	0.102
問2-2⑤ 訪問調査では、機構の評価担当者との間で、教育研究活動等の状況に関する共通理解を得ることができた	0.656	0.038	0.035
問2-2⑦ 訪問調査時の機構の評価担当者は十分に研修を受けていたと思う	0.605	0.016	0.087
問2-2⑥ 訪問調査時の機構の評価担当者の人数や構成は適切であった	0.552	0.148	-0.076
問2-1④ 貴校の総合的な状況が広く社会等の理解を得るために、わかりやすい自己評価書にすることができた	0.014	0.864	-0.036
問2-1⑤ 自己評価書の完成度は満足できるものであった	-0.043	0.772	-0.015
問2-1① 評価基準および観点に基づき、適切に自己評価を行うことができた	0.174	0.413	0.175
問2-2① 訪問調査の前に提示された、「書面調査による分析状況」の内容は適切であった	0.022	-0.025	0.830
問2-2② 訪問調査の前に提示された、「訪問調査時の確認事項」の内容は適切であった	-0.019	0.015	0.785
因子間相関		0.474	0.629
			0.417

別表 A-3 説明会・研修会に対する因子分析結果

(N=122)

質問項目	第1因子	第2因子
	説明・研修会 の理解	機構の有効な 対応
問4① 説明会の配付資料は理解しやすかった	0.908	-0.039
問4⑤ 自己評価担当者等に対する研修会の内容は理解しやすかった	0.907	-0.038
問4② 説明会の説明内容は理解しやすかった	0.902	-0.099
問4④ 自己評価担当者等に対する研修会の配付資料は理解しやすかった	0.771	0.068
問4③ 説明会の内容は役立った	0.635	0.119
問4⑨ 機構事務局の対応（質問等に対する対応）は適切であった	-0.061	0.706
問4⑧ 機構の教職員が行った訪問説明は役立った	-0.023	0.693
問4⑦ 機構が配布している自己評価実施要項等の冊子は役立った	0.346	0.436
因子間相関	第1因子	0.530

別表 A-4 機構の評価結果（評価報告書）に対する因子分析結果（N=129）

質問項目	評価報告書の満足
問5-1⑨ 総じて、機構による評価報告書の内容は適切であった	0.821
問5-1④ 評価報告書の内容は、貴校の目的に照らし適切なものであった	0.811
問5-1⑧ 評価報告書の構成および内容は分かりやすいものであった	0.773
問5-1② 評価報告書の内容は、貴校の教育研究活動等の改善に役立つものであった	0.765
問5-1⑥ 評価報告書の内容は、貴校の規模等（資源・制度など）を考慮したものであった	0.764
問5-1⑤ 評価報告書の内容は、貴校の実態に即したものであった	0.764
問5-1① 評価報告書の内容は、貴校の教育研究活動等の質の保証をするために十分なものであった	0.730
問5-1③ 評価報告書の内容は、貴校の教育研究活動等について社会（学生・保護者、企業、その他関係者など）の理解と支持を得ることを支援・促進するものであった	0.632
問5-1⑦ 評価報告書の内容から、教育研究活動等に関して新たな視点が得られた	0.507

別表 A-5 自己評価による効果・影響に対する因子分析結果（N=129）

質問項目	第1因子	第2因子
	意識変化	現状把握および改善促進
問6-1③ 教育研究活動等を組織的に運営することの重要性が教職員に浸透した	0.996	-0.132
問6-1⑨ 自己評価を行うこと的重要性が教職員に浸透した	0.718	0.041
問6-1④ 各教員の教育や研究に取り組む意識が向上した	0.675	0.164
問6-1① 教育研究活動等について全般的に把握することができた	-0.062	0.830
問6-1② 教育研究活動等の今後の課題を把握することができた	-0.070	0.691
問6-1⑤ 貴校の教育研究活動等の改善を促進した	0.233	0.491
問6-1⑧ 貴校の個性的な取組を促進した	0.228	0.479
因子間相関 第1因子		0.612

別表 A-6 機構の評価結果による効果・影響に対する因子分析結果（N=128）

質問項目	第1因子	第2因子	第3因子	第4因子
	意識変化	現状把握	理解と支持	改善促進
問6-2③ 教育研究活動等を組織的に運営することの重要性が教職員に浸透する	0.936	0.028	-0.153	0.009
問6-2⑨ 自己評価を行うこと的重要性が教職員に浸透する	0.879	-0.099	0.038	0.014
問6-2④ 各教員の教育や研究に取り組む意識が向上する	0.540	0.110	0.025	0.202
問6-2⑩ 教職員に評価報告書の内容が浸透する	0.451	0.008	0.317	0.017
問6-2① 教育研究活動等について全般的に把握することができる	-0.038	0.889	0.022	0.002
問6-2② 教育研究活動等の今後の課題を把握することができる	-0.041	0.878	-0.091	0.044
問6-2⑫ 貴校の教育研究活動等の質が保証される	0.247	0.408	0.177	-0.152
問6-2⑬ 学生（今後入学する学生を含む）の理解と支持が得られる	0.035	-0.056	0.922	-0.056
問6-2⑭ 広く社会の理解と支持が得られる	-0.109	0.029	0.827	0.075
問6-2⑦ 貴校のマネジメントの改善を促進する	0.078	-0.080	-0.076	0.820
問6-2⑤ 貴校の教育研究活動等の改善を促進する	0.096	0.108	0.018	0.588
問6-2⑧ 貴校の個性的な取組を促進する	-0.033	0.059	0.214	0.514
因子間相関 第1因子		0.541	0.540	0.718
第2因子			0.483	0.623
第3因子				0.534

[ABSTRACT]

**Model of Institutional Quality Enhancement Promoted by “Certified Evaluation and Accreditation”
: Comparative Analysis among Different Types of Higher Education Institutions**

KIM SoungHee * , HAYASHI Takayuki ** , SAITO Takahiro **

This research examined the effectiveness of “Certified Evaluation and Accreditation” conducted by the National Institution for Academic Degrees and University Evaluation (NIAD-UE). Statistical analysis of responses to a questionnaire survey of evaluated institutions produced the following findings: (1) the evaluation of Colleges of Technology affected their improvement because of a better consensus on evaluation standards among external evaluators and institutions, and (2) a strict accreditation process for law schools caused some conflicts among evaluators and institutions. A structural equation model confirmed that an appropriate self-evaluation of current education and research activities helps to promote quality improvement within an institution, and the evaluation result works as an external pressure and/or incentive for quality improvement. NIAD-UE needs to respond to these counteractions concerning the evaluation process and results inside institutions to develop a more effective evaluation system as a whole.

* Doctoral Candidate, Graduate School of Decision Science and Technology, Tokyo Institute of Technology

** Associate Professor, Department of Research for University Evaluation, National Institution for Academic Degrees and University Evaluation