

【研究ノート】

多様化する学士課程教育の自己評価方法

串本 剛

広島大学大学院教育学研究科・日本学術振興会特別研究員

t.kushimoto@hiroshima-u.ac.jp

要 約

本研究の目的は、評価方法に関する4つの類型を設定し、学士課程教育の自己評価におけるそれらの採否状況について、全国調査から得られたデータを用いて実証的に明らかにすることである。前半部分では、学士課程に関する教育観の違いを反映するものとして、自律型、応需型、標準化型、成果準拠型という4つの評価方法を導入する。後半では、全国1,871の学部を対象として実施した質問紙調査の結果を使って、各評価方法の採用状況を実証する。評価方法の採用率は、評価領域や学部の属性により異なることが明らかになると共に、その違いが評価方法の背後にある教育観との関係から考察される。

キーワード

学士課程教育、自己評価、評価方法、多様化、全国調査

1. 研究の背景と目的

(1) 研究の背景

制度的かつ恒常的な取り組みとしての評価は、近年、産官学を問わず広く関心の対象となってきた（牟田 2001）。「学」の一部をなす高等教育も例外ではなく、最近の動向で言えば、すべての大学が7年以内ごとに文部科学省の認証を受けた評価機関による評価を受けなくてはならないとする認証評価制度が、2004年から導入された。

現在、学士課程教育を対象とする認証評価機関には、大学基準協会、大学評価・学位授与機構、日本高等教育評価機構の3つがあり、評価基準や項目、費用等の面では異なる枠組みを利用してい

る¹。ただし共通点もあり、そのひとつが大学の実施する自己評価に基づいて評価を行うという点である。自己評価を基礎とする第三者評価という仕組みは、日本に限らず世界的にも見られるものであり、それだけ大学にとって組織の自律性が重要であることを示していると言える。特に日本では、自己評価の制度化（1991年）の後、その一定の普及を経てから外部評価の努力義務化（1999年） 認証評価制度の導入と展開してきており²、今日でも大学をめぐる評価において自己評価が中心に位置づけられていると見ることができる。

自己評価とは、評価のプロセスにおける価値判断の主体がプログラムの実施者である、ということの意味するに過ぎないが、学士課程教育を対象としたそれは従来、価値判断だけ無くその際に考

慮される情報（「根拠情報」）すらも、主として評価対象に対する「自己」の主観的判断に負っていた³。例えば教育内容について評価しようとするとき、それについて自分たちがどう思うかということが、主に考慮されてきたのである。

しかし、各種の大学改革や上記の第三者評価の進行、あるいは大学を取り巻く社会経済的環境の変化に伴い、いまでは授業評価等を通して学生の意見が顧慮されるだけでなく、外部評価によって学外の主体による判断を参考にする機会も増えている。さらに教育の目的が達成されているかという根本的な問いに回帰し、教育成果に関する客観的・主観的な情報を把握し、それらに基づいて教育の是非を判断するという動きも見られる。つまり、評価方法の多様化が起きているのである。

（2）本論の目的と構成

こうした状況の推移が、「何を持って教育の是非を判断すべきか」という教育観の多様化とも並行するものであるとすれば、自己評価の実態解明には、単なる現状把握以上の意義があると考えられる。だが第三者評価のそれに比べ、大学の自己評価に関する研究蓄積は限られており、根拠情報との関連からその評価方法を明らかにするという試みは、世界的に見てもこれまでのところ為されていない⁴。

そこで本稿では、幾つかの異なる教育観を反映するものとして評価方法の類型を設定し、学士課程教育の自己評価におけるそれらの採用状況について、全国調査から得られたデータを用いて実証的に明らかにすることを目的とする。この結果得られる知見には、これまでブラック・ボックスであった現象に光を当てることだけでなく、変革期にある学士課程教育に関する議論にも、ひとつの視点を提供することで一定の貢献をすることが期待できる。

以下、第2節ではまず、分析の視点となる4つの評価方法の相互関係と、それぞれが依拠する教育観、および評価方法採用の徴拠となる根拠情報について解説する。続く第3節では、2006年に実施した全国調査の結果を使って、評価方法の採否状況について検討する。分析に際しては、評価領域

による相違と共に、学部の属性と採否の関係についても検証する。最後に第4節では、学士課程に関する教育観の視点から分析結果を考察し、今後の課題を提示する。

2. 分析の視点～4つの評価方法

（1）4類型の関係

評価の類型化には、ふたつの方法がある。ひとつは、「評価とは何か」という評価の定義の違いをも反映させる形で行われるもので、この場合、各類型はある程度背反的であることが前提とされており、「モデル」と称されることが多い。（プログラム）評価論一般の研究では、こうした類型化がたびたび試みられており、近年のものとしては、22のモデル（アプローチ）を4つのカテゴリーにまとめて論じたStufflebeam（2001）や、哲学的（客観主義 - 主観主義）方法論的（量的 - 質的）、実践的（情報提供のみか価値判断を含むか、評価者が指導的な役割を果たすか援助的か、評価者の経験、評価対象に関する専門知識が必要か否か、依拠するモデルがひとつか状況に応じて選択的か）な相違を考慮した上でHouse（1983）を参考にしつつ、5つのアプローチを示したFitzpatrick et al.（2004）がある。

アメリカの先行研究には、高等教育を念頭に置いたものの中にも、プログラム評価論の知見を援用してこのような類型化を行っているものが見られる。Gardner（1977）は初期の代表的な研究で、評価の時系列的な進化に対応させて5つの評価モデルを提示している。また説明責任の遂行に加え資源配分の根拠として評価が利用されるようになった頃のConrad and Wilson（1985）では、Gardner（1977）を踏襲する形で、目標標準拠（objective-based）、応答的（responsive）、意思決定（decision-making）、鑑識眼（connoisseurship）という4つの評価モデルを挙げている。

これらに対し、高等教育に関する日本の先行研究では、何らかの評価の定義をしばしば暗黙の前提として、評価に関わる諸要素について類型化を試みる傾向にある。例えば関（1989）は、評

価の性格、時期、過程・対象、主体、機能、方法の各側面から、同様に大学基準協会(2000)では、評価の主体、対象、内容、時期、対象の動態、結果の効果という切り口から類型化を行っている。この方法の場合、類型間の排他性は前提とされておらず、ある評価が複数の類型に当てはまるということもあり得る。どちらの方法による類型化が有効かは、研究の目的や関心により異なり、カリキュラムに関する著作で評価モデルを論じている Stark and Lattuca (1997) のように、上記の2つの中間的な方法による類型化の例もある。

研究の背景として述べた通り、本稿の関心はあくまでも学士課程教育の自己評価にあり、その評価方法の多様性についても、完全な移行ではなく、伝統的なそれに対する追加的な導入を前提としている。従って、本論における類型化は、どちらかといえば後者の方法によるものであり、4つの評価方法の類型(アプローチ)は、基本的に並存するものとして提示される。

(2) 自律型

最初の評価方法は、学士課程教育が「所属教員の専門性や見識を尊重し、自律的に営まれるべきである」という教育観を前提とするもので、「自律型」と名づけることができる。この評価方法では、根拠情報として評価対象に対する所属教員の主観的な判断が用いられるため、場合によっては評価のための情報収集が組織的に行われず、ということもあり得る。

自律型アプローチは、最も伝統的な評価方法であり、自己評価が制度的に位置づけられる以前から採用されてきたものと言える。なぜならば、その背景には中世以来の大学の自治、そして19世紀にドイツで興った近代大学における学問の自由という理念が存在するからである。古くは王権や宗教への従属、近代以降で言えば国家による統制や産業界への過度な迎合は、それらの理念を脅かすものとされてきたため、大学には外界からの干渉を忌諱してきたという歴史的経緯がある。これが、学士課程教育の是非はその教員によって判断されるべきだという、教育観に結びついているものと考えられる。

こうした発想は、Eisner (2003) の言う鑑識と批評 (connoisseurship and criticism) に依拠しているものと考えられ、従来の類型化でもしばしば取り上げられてきた (Gardner 1977, Conrad and Wilson 1985)。通例Eisnerのモデルは、アメリカにおけるアクレディテーションのような外部評価を説明する概念として用いられてきたが、より日常的な現象に目を向ければ、多くの自己評価がその具体例だということがわかる。

また、常にプログラム全体として評価が行われるわけではなく、各教員の個別授業などにおける日常的な評価の積み重ねが、組織的な行為を代替し得るという側面に注目すると、Stark and Lattuca (1997) が挙げている「非公式の評価 (informal assessment)」に該当するとも言える。

自律型の評価方法は、学士課程教育という評価対象の専門性の高さを考慮した場合、有力なアプローチであることは間違いない。ただし Fitzpatrick et al. (2004) が指摘するように、非公式の評価は評価者の経験や臆見により偏向してしまう危険性がある。

さらに、専門性に依拠した評価自体の妥当性が問われる際には、正当化が難しいことも事実である。これは単に、もはや象牙の塔ではなくなったという社会的な位置づけの変移だけではなく、教員個人の省察は教育観の制約を受けるため改善につながるとは限らない (Trigwell and Prosser 1996) というような、学問的知見の広まりにも由来した意識転換の結果でもある。こうした価値観の動揺が、以下に挙げる3つのアプローチの台頭を促していると見られる。

(3) 応需型

応需型の評価方法は、「学生という「第一の顧客」の要求に応えるべきである」という教育観に根ざすもので、評価対象に対する学生の意見を根拠情報として用いる。

「第一の顧客」とは、非営利組織について論じたドラッカーとスターン (訳書 2000) の用語で、「ある活動によって生活が改善される人々 (p. 20) のことを指す。一般に何らかの行為 (事業プログラム) には、それによって益すると考えられる

「顧客」が複数考えられる。学士課程教育も例外ではなく、学生以外にも、その保護者や卒業生の雇用者、さらに活動を通して達成感を得るという意味では、教員や職員もそれに含まれる。だが、教育を行う一義的な意図が彼/彼女らの変容にあるという意味では、学生こそが「第一の顧客」と言える。

営利的な活動であれば、応需型アプローチは基本的な評価方法とさえ言えるものであろう。しかし大学においては、専門知識・技能を持たない者（学生）がその所有者（教員）からそれらを教授されるという関係性が支配的であったため、学生の意見が評価の根拠となるという発想はなじみにくかった。その名残は、授業評価が導入される時に繰り返される「学生に評価をする能力（資格）はない！」というお決まりの台詞からも窺うことができる。セイモア（訳書 2000）は、Garvin（1988）による質の定義を引きながら、こうした問題を鮮明に解説している。

しかし学生が権利の主体、それも教育というサービスを購入する消費者として認識されるようになるのに伴い、「第一の顧客」である学生の意見を考慮することの必要性が認められるようになってきている。学生が消費者として位置づけられるに至った経緯については、1980年代初頭にリースマン（訳書 1986）がアメリカの文脈において論じているが、最も大きな要因は大学が大衆化することにより、学生の関心や能力が多様になったことである。そうした傾向は今日では多くの国々で見られ、学生は権利と具体的な期待・要望を持っているのだから顧客である、という認識は広く認められている（Davis 2002）。実際にわが国でも、「学生を、大学にとって主たるクライアントであると位置づけ直し、そのニーズを汲み取り、そのうえで教育プランを設定する」（日本私立大学連盟 学生部会 2000、p. 203）ことの必要性が訴えられるようになってきている。

事実、「第一の顧客」の意見を根拠情報とすることに特徴を求める評価モデルは、比較的新しい研究にしか見られない。Stark and Lattuca（1997）では、これを学生中心評価（student-centered evaluation）と呼び、「様々な目標を持った学生の授業やプログラムでの経験を知ることにより、教

員と学生の目的意識の一致度を調べる」（p. 291）作業であると位置づけている。本研究で言う応需型アプローチは、必ずしもそのような評価目的を前提とするものではないが、学生の意見を考慮するという点では、学生中心評価に共通するものである。

（4）標準化型

「機関を超えて、同等の教育を提供すべきである」という教育観が、第3の評価方法である「標準化型」を支持するものである。ここではプログラムの外部から得られる情報が考慮されるため、教育が標準的な形に収斂していく可能性が内包されている。

同等性や通用性に価値を置く教育観は、大学が学位という世界共通の資格を発行する唯一の機関であるとされることからわかるように、大学の根本的な理念のひとつであると言える。ただし、その担保の仕方は時代と共に変遷してきており、概括的に言えば、かつては大学あるいは課程の設置時に同等性を保証するための仕組みが設けられていたのに対し、近年では事後的にそれを行う傾向にあり、標準化型アプローチによる自己評価はその機能を担っているということが出来る。こうした移行は、高等教育セクターの肥大化の結果、行政の役割が変化することによって起こるものであり、設置基準の大綱化に伴う自己評価の努力義務化という事実によって、日本においては顕現している。

「評価対象の標準化に対する志向」という方法上の特徴は、従前の評価モデルには見られないもので、ある意味で学士課程教育の自己評価独特の方法と言える。しかし、具体的な根拠情報を考えると、例えば学外の専門家による見解を考慮する場合は、自律型アプローチのところ而言及したEisnerのモデルを応用しているものと捉えられる。このとき注意すべきは、学外の専門家にはその専門性もさることながら、広く事情に精通していることによる見解の普遍性が期待されていることである。換言すれば、彼らの見解を参照することにより、当該教育の通用性が保証されることになる。こうした力点の違いが、ひとつの評

価モデルの原理が、ふたつの評価方法の文脈で参照される原因となっている。

標準化型アプローチを支える教育観が、学士課程教育にとって本質的なものである反面、高等教育の大衆化によって、学士課程教育が多様な形態・内容を持たざるを得なくなる状況の中で、一元性を志向するその評価方法自体に限界があることも確かである。次項で取り上げる成果準拠型が必要となる背景には、そうした事情も存在すると考えられる。

(5) 成果準拠型

ここまで論じてきた3つの評価方法とは異なり、行為の是非をその結果によって判断しようとする「成果準拠型」アプローチは、「教育の妥当性はその成果によって判断されるべきである」という教育観から発生する。そのため根拠情報としては、何らかの指標によって把握された教育成果が求められる。

教育とは、教育者の意図を学習者のうちに実現することを目的とする行為であり、その意味で成果準拠型アプローチが前提とする教育観は、至って常識的なものであると言える。しかし応需型の背景のところでも論じたように、大学における教員と学生の関係は、必ずしも上記の意味での「教育的関係」ではなかったため、教育目的や目標については教育成果を明確にするということは、一般的ではなかった。

こうした状況が変わりつつあるのには、主としてふたつの原因が考えられる。ひとつは、1980年代以降世界的に見られる説明責任論の隆盛の中で、結果に対する関心が高まっているということである。そこでの焦点は、対外的に教育成果を示すことにあると考えられるが、それが結果として自己評価における教育成果の重視にも繋がっている。もうひとつ、構成主義的な学習理論が大学教育の文脈でも重視されるようになってきたため、教育そのものの質を見ることによりその成果を推認することの妥当性が問われている、ということも見逃せない。この点については、Barr and Tagg (1995) 等が大学教育のパラダイム転換として論じているほか、評価に関連してはHuba and Freed

(2000) やPalomba and Banta eds. (2001) が言及している。

成果準拠型アプローチは、Conrad and Wilson (1985) によると体系的に論じられているものとしては最も古い評価モデルである、目標準拠モデル (objective-based model) の萌芽的な形態と見ることができる。その祖がTyler (1949) に求められる目標準拠モデルでは、教育目標を操作的に定義することにより達成の測定を可能にし、目標の達成度をもって教育の質を判断するという構造が前提とされる。本論における成果準拠型アプローチは、ここまで体系的な構造を有するわけではないが、教育の評価に際しその成果を考慮する点において共通の特徴がある。

教育成果を根拠情報として用いることには、原理的な正当性を認められるだけでなく、教育そのものの多元性を正当化しようという利点もある⁵。だが、実用の難しさもしばしば指摘されるところであり、その代表的なものは、教育成果が把握されただけでは何を改善すべきかがわからない、ということである。わが国でも教育成果への関心が高まる中、同様の批判が見られる (大山 2001)。しかしこの問題に関しては、プログラム・セオリーやロジック・モデルと呼ばれる、原因と結果の関係性を明確化することによる対応策も論じられてきており (例えばRogers et al. 2000) 教育を対象とした評価においても徐々に浸透してきている⁶。

3. 評価方法の採否

(1) 分析データ

前節で提示した4つの評価方法の採用状況を明らかにするために本論で用いるデータは、学士課程教育の自己評価の実態を解明することを目的として開発した「教育評価と教育改善に関する調査」の調査結果である。

調査票は、学部の特徴、教育評価体制等、教育評価の方法、教育評価の効果・影響、教育評価に関する回答者の意見、について聞く5つの部分から構成されている。教育評価の体制に関しては経年比較の可能性を考慮し広島大学・大学教育研究

センター編（1999）の調査を、効果・影響については適切な日本の先行研究がなかったため、アメリカで実施された類似調査である Institutional Support for Student Assessment Inventory（ISSA）を参考にした⁷。また本論の主題である評価方法に関する問は、串本（2006）でまとめている自己評価報告書の分析や、第三者評価機関の取組を基に設定した。

調査対象は2005年5月の時点で学士課程を持つ全国1,871の学部で、質問紙は4月21日を締め切りとして、2006年3月28日に発送された。4月末には督促はがきを送り、最終的に6月15日時点での回収率は31%であった（表1）⁸。

表1 調査の概要

	国立	公立	私立	全体
発送数	345	158	1,358	1,871
回収数	215	53	309	580
回収率	62.3%	33.5%	22.8%	31.0%

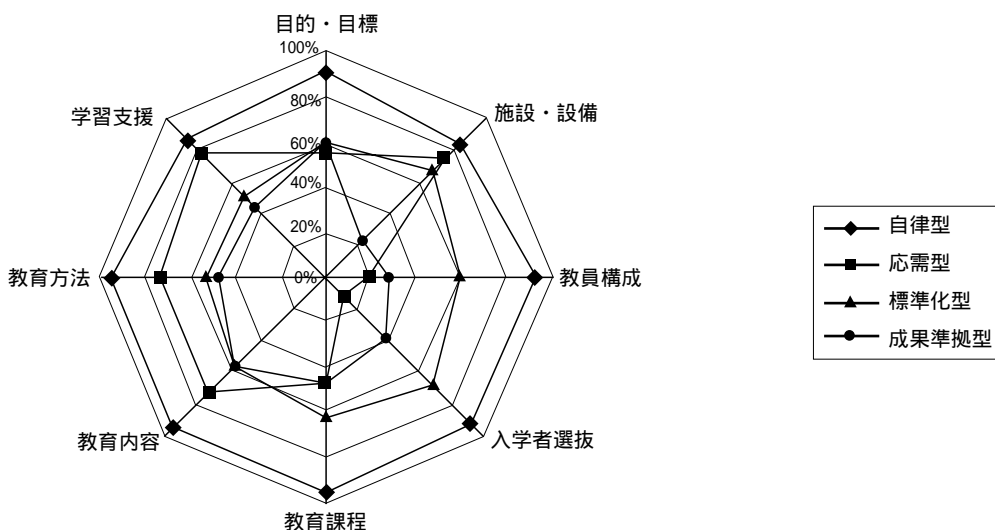
（出所）筆者作成

（2）全体的傾向

評価方法に関する問いでは、下記1）～8）の各領域について、現状の適否や改革の必要性を判断する際にどのような根拠情報を用いているかを、複数回答で尋ねた。その回答に基づき、各領域における評価方法の採用状況を示したものが図1である⁹。

- 1) 目的・目標：教育の理念・目的・目標の内容・具体性など
- 2) 施設・設備：教室、実験器具、情報インフラ、図書館など
- 3) 教員構成：職階・年齢・専門分野の構成、教育者としての素養など
- 4) 入学者選抜：入試の種類・内容など
- 5) 教育課程：コース・課程の設定、各種科目の配列・割合など
- 6) 教育内容：提供科目、授業内容など
- 7) 教育方法：授業方法、授業形態・規模、シラバス、成績評価方法など
- 8) 学習支援：履修指導（ガイダンス、オフィスアワー）、補習体制など

図1 評価方法の採用状況



（出所）筆者作成

自律型はいずれの領域でも8割以上の学部が採用しており、最も広く用いられている評価方法であることが確認できる。既述の通り評価方法は併用を前提としているので、むしろ自律型アプローチを「使わない」という事例は稀であると解釈した方がわかりやすい。

評価領域間で採用率に差がないという点では、標準化型の評価方法も共通している。採用率は全体に6割前後であり、普及が十分に進んでいるとは言い難いが、外部から獲得された情報は領域を問わず、評価の際に考慮されていることがわかる。

上記のふたつとは対照的に、応需型アプローチの採用状況には、評価領域によって格差がある。教員構成や入学者選抜の評価ではほとんど利用されていないのに対し、その他の6つの領域ではほぼ半数以上の学部が採用している。

成果準拠型の評価方法も、応需型に似て領域間で採用程度に差がある。また、採用率は全体に低く、最も普及している目的・目標の評価においても、それは6割に満たない。その成果を根拠として教育条件や教育活動の是非を判断するという方法は、大学においては依然として定着していないと言える。

全体的な傾向からは、自律型を基礎にしつつも評価方法の多様化が進んでいることが実証されたが、その様相は評価領域によってかなり異なることも明らかとなった。では、評価方法の採否には、学部の属性も関係するだろうか。次項でその点を論ずる。

(3) 評価方法の採否と学部の属性

評価方法の採否と関係があると考えられる学部属性として、ここでは設置形態、専門分野、卒業生の職業に関する想定の方針、入試難度の4つを設定した。専門分野は、学習における自由度の高さを反映すると考えられる文系・理系の別により、職業想定の方針は関連する問いの回答に基づいて、また入試難度は学部偏差値に従って下位と上位に区別した¹⁰。

表2には、評価領域毎に行った評価方法の採否と学部属性のクロス集計結果を、4つの評価方法についてまとめている。表中の値はすべて当該ア

プローチを採用している学部の割合で、*印はカイニ乗値による検定結果である。

まず自律型について見ると、いずれの領域でもほとんどの学部が採用していることもあり、有意差が見られる関係は多くない。目的・目標、教員構成、学習支援では設置形態、施設・設備においては専門分野や職業想定程度の違いで有意差が見られる。

応需型アプローチでは、有意差が見られるのは教育課程の評価についてのみだが、その他5つの領域でも理系より文系の方が採用率が高くなっている。専門分野以外を見ると、教育方法と学習支援は設置形態の違いで、施設・設備は職業とのつながりの強弱で採用率に開きが見られる。特に設置形態に関しては、教育課程を除くすべての領域で公立大学での採用率が低い。

標準化型の採否と学部属性の間には、一転して統計的に有意な関係が数多く確認できる。まず設置形態では、私立大学における採用率の低さが顕著であり、応需型の場合とは対照的に、公立大学で採用率が高くなっていることも特筆に値する。専門分野による違いも明らかで、ほぼすべての領域で理系学部の採用率が高くなっている。

最後の成果準拠型は、標準化型以上に職業との繋がりによる差異が大きく、設置形態や文理の別では有意差が見られない。これに加え、入学者選抜、教育内容、教育方法では入試難度による採用率の相違も見られる。

以上の分析からは、自律型や応需型に比べ、標準化型や成果準拠型では学部属性との間に顕著な連関が見られ、特に前者では設置形態や専門分野、後者では職業との繋がりという属性が、違いを鮮明に描き出したと指摘できる。

4. 考察と今後の課題

(1) 領域間格差について

ここまで示してきた分析結果に対し、各評価方法が前提としている教育観の視点から考察を加えると、幾つかの知見を提供することができる。

各評価方法の全般的な採用率からは、自律型を

表2 評価方法の採否と学部属性

(%)

	全体	設置形態			専門分野		職業想定		入試難度	
		国立	公立	私立	理系	文系	否定	肯定	下位	上位
自律型										
目的・目標	91.2	95.1	90.9	88.5*	92.0	91.1	94.0	90.4	92.9	90.1
施設・設備	84.0	88.2	84.8	80.9	89.1	81.6*	76.1	86.1*	82.5	86.1
教員構成	92.7	96.5	95.5	89.5*	93.2	92.2	89.2	93.6	92.4	92.8
入学者選抜	91.8	94.9	93.5	89.2	92.3	91.5	89.8	92.3	90.0	93.1
教育課程	95.5	96.5	100.0	94.0	94.4	96.2	97.4	95.2	95.5	95.1
教育内容	95.1	96.5	91.3	94.8	94.4	95.6	92.3	95.9	95.5	95.2
教育方法	95.0	95.5	91.3	95.1	94.9	94.9	94.0	95.2	95.5	94.0
学習支援	86.5	92.4	84.4	82.6**	86.2	86.3	86.4	86.7	83.8	88.6
心需型										
目的・目標	55.1	54.7	43.2	57.0	50.8	58.3	51.3	56.1	55.4	55.3
施設・設備	73.2	77.9	60.9	72.2	72.5	74.7	60.6	76.8**	73.1	74.4
教員構成	19.8	19.7	15.9	20.7	15.6	21.8	18.9	20.1	19.1	20.4
入学者選抜	12.3	14.6	10.9	10.8	12.4	12.4	16.7	11.1	11.7	13.4
教育課程	47.2	44.3	47.7	49.3	39.5	52.2**	47.4	47.3	50.2	45.7
教育内容	72.2	76.1	58.7	71.5	69.2	74.4	68.4	73.1	71.1	73.7
教育方法	73.8	80.6	67.4	70.0*	75.8	73.3	68.4	75.2	73.7	74.6
学習支援	77.7	84.8	71.1	73.7**	80.6	76.5	71.8	79.4	77.1	78.4
標準化型										
目的・目標	57.3	66.5	63.6	49.7**	65.8	52.9**	52.1	59.0	56.7	58.6
施設・設備	66.1	68.7	65.2	64.6	71.0	64.1	63.3	67.0	65.8	66.5
教員構成	59.0	67.2	70.5	51.1**	65.1	55.7*	51.4	60.9	56.4	61.9
入学者選抜	67.2	74.7	71.7	61.0***	67.0	68.3	63.0	68.3	65.0	70.6
教育課程	61.9	69.7	72.7	54.6**	67.7	58.6*	53.5	64.5*	56.7	67.2*
教育内容	55.8	63.2	69.6	48.3**	65.7	49.5***	45.3	59.0**	52.4	58.9
教育方法	52.5	56.7	54.3	49.5	62.6	47.6**	42.7	55.3*	47.8	58.6*
学習支援	51.0	53.8	60.0	47.7	53.1	50.7	42.7	53.4*	49.6	53.4
成果準拠型										
目的・目標	59.0	60.6	61.4	57.3	63.8	56.1	46.2	62.9**	57.9	60.8
施設・設備	22.7	24.6	21.7	21.7	25.9	21.1	13.8	25.2*	24.4	21.4
教員構成	27.9	27.8	25.0	28.6	27.6	28.0	23.4	29.2	25.8	28.7
入学者選抜	38.5	39.9	39.1	37.5	43.8	35.3	25.9	41.9**	34.2	43.1*
教育課程	47.5	47.3	52.3	47.2	50.3	46.2	32.5	51.7***	47.3	49.4
教育内容	56.7	61.7	50.0	54.2	60.6	54.3	41.0	61.2***	50.4	62.6**
教育方法	46.8	52.2	47.8	42.7	51.3	44.1	35.0	50.4**	42.1	52.8**
学習支援	43.1	43.7	44.1	42.3	45.4	42.2	28.2	47.3***	40.0	46.2

(出所)筆者作成

(注)*** p<0.001, ** p<0.01, * p<0.05

基礎としながら、他のアプローチが導入されることにより評価方法の多様化が進んでいるという実相が示された。8つあるいずれの評価領域でも、自律型アプローチの採用率が最も高く、教育の是非を判断する際には所属教員の専門性や見識を尊重すべきであるという発想が、確実に存在していると言える。

評価方法の多様化は、そうした発想の再検討を迫るものとして捉えられるが、必ずしもすべての評価領域において、それが均一に起こっているわけではない。標準化型アプローチの採用に表れる同等性・通用性を重視する傾向が領域を問わず見られる一方で、幾つかの領域で利用率が3割にも満たない評価方法も存在する。

応需型の評価方法は、教員構成や入学者選抜方法の評価では採用率が低く、そこでは学生の要求に応えるべきだという考え方が浸透していないと言える。入学者選抜方法の評価では、現在の学生ではなく未来の学生つまり高校生の意向を取り入れているという見方も可能だが、教員構成に関しては、学生にその良否を尋ねることは稀であると結論できるだろう。教員の職務は教育だけではないことを考慮すると当然の結果ともいえるが、教育内容や教育方法の評価では多くの学部が学生の意見に配慮しており、この隔たりを埋めるための取り組みが、近く義務化されるFD活動なのかもしれない。

もう一点、採用率の低さが顕著なのは、施設・設備や教員構成の評価における成果準拠型アプローチである。これらの領域は教育成果との繋がりを確定しづらいため、それによって是非を判断するという着想に至らないものと考えられる。

(2) 学部の属性による格差について

評価領域による格差に加え、学部の属性によっても評価方法の採用率に違いが見られた。これに関しても教育観の視点から説明できる点が幾つかある。

学部の属性で採用率に有意差がよく表れたのは、標準化型と成果準拠型の評価方法であり、前者に関しては、特に設置形態による相違が目立った。すべての領域で、私立大学での採用率が低い

のに対し、個性化の必要性が訴えられる今日においても、公設大学ではある程度の同等性が志向されていると言える。

また教育課程や教育内容の評価に関して公立大学で採用率が高くなっているのは、卒業生の職業を想定している学部で同様の傾向が見られることから、農学系・保健系の学部の多さが原因になっていると推測できる。これらの分野 農学部には獣医学科がある例も多い には、国家試験という出口の標準化を進める仕組みが備わっているため、教育そのものの評価においても標準化型が選好されると考えれば、専門分野では理系の、入試難度では上位の学部において採用率が高くなっていることも納得できる。

成果準拠型アプローチは、専門分野や設置形態でははっきりとした差が見られず、卒業生の職業が具体的に想定されている学部における採用率の高さが顕著であった。これは、就職状況や資格の取得により適切な教育成果が定義しやすいことが関係していると考えられる。領域間格差の原因と総合すると、成果を示しやすく、成果と教育との因果関係が確定しやすい場合に、より多く採用されていることになる。

なお、応需型の評価方法に関しては、学部の属性による格差はあまり見られなかったが、特に学生の意見を把握する必要があるのはどのような場合か、ということを考えて、入試難度による差異が期待された。なぜならば、葛城(2006)も示しているように、入試難度が下位の大学に在籍する学生ほど、従来想定されていた学生像とは異なることが認められているからである。学生の変質が遍く起きている証左ともとれるし、あるいは、「意見を聞くに値する学生」という教員の意識が入試難度上位の大学で働き、仮説に従う傾向を相殺している可能性もある。

(3) 今後の研究課題

今後着手が求められる研究課題としては、評価方法の採用状況に関する時系列変化の解明が挙げられる。

今回の分析からは、評価方法の多様化の実態が明らかとなったが、これは大学改革や第三者評価

の影響下において急変している過渡期のそれを描出しているとも言える。従って、評価方法の選択は、必ずしも背後にある教育観を了承した上で、意識的に行われているとは限らない。

この事実は裏を返せば、評価領域や学部属性による採用率の格差を、意図的に是正する余地が存在するという点でもある。例えば、「教育の妥当性はその成果によって判断されるべきである」という教育観は、現状では成果の定義がしやすい学部において配慮される傾向にあるが、5年10年経った後には、学部属性を問わず重要性が認識され、成果準拠型アプローチが普及しているかもしれない。

時系列的变化を追うことにより、単にその時々の実態が判明するだけでなく、教育観の変遷を知るための一指標を見出すこともできると考えられる。

注記

- 1 各機関が用いる評価システムの枠組みや評価結果については、ホームページから知ることができる。アドレスはそれぞれ、大学基準協会 (<http://www.juaa.or.jp/>)、大学評価・学位授与機構 (<http://www.niad.ac.jp/>)、高等教育評価機構 (<http://www.jihe.or.jp/>)。
- 2 ここでは紙幅の関係から割愛するが、詳しくは清水 (2004) 等で論じられている。
- 3 評価の定義は論者により多様だが、本稿では、「指定および把握した情報を何らかの基準に照らすことにより、ある対象の価値を判断すること」という比較的汎用性の高い定義を採用している。また本文で述べている「根拠情報」とは、この定義に含まれる「情報」を指し、評価対象に関する情報だけでなく、評価の際に考慮される情報全般を含む。従って上記の定義は、他者の評価を参考に自己評価する、というような「評価の入子構造」を許容するものである。
- 4 自己評価の方法に関心を払った日本の先行研究としては、広島大学・大学教育研究センター（現高等教育研究開発センター）が1999年に実施した全国調査をもとに評価項目分析を行った米澤他 (2000) が挙げられる。ただしここでは、評価項目に注意が向けられており、根拠情報の質的な違いは配慮されていない。またアメリカの類似研究である Peterson and Augustine (1999) では、Student Assessment (SA) と学事に関する意思決定の関係について論じているが、根拠情報のひとつであるSAの結果を相対的に位置づけることには成功していない。
- 5 技術者教育のプログラム認定機関であるJABEEが採用する評価の枠組みは、成果準拠型アプローチのこうした特徴を念頭に置いた実例のひとつである。詳しくは、<http://www.jabee.org/OpenHomePage/q&a0204-0509.htm>参照。
- 6 初等教育での実践例に関しては串本・渡辺 (2006) を、大学教育の文脈での議論としては串本 (2006) を参照されたい。
- 7 ISSAは自己評価全般に関する調査ではなく、学生についての情報把握 (student assessment) を対象としており、本論の関心と厳密に一致するわけではない。しかし全米規模の当該テーマに関する画期的調査であり、質問項目も参考になるものが少なくない。同調査については Peterson et al. (1999) を参照されたい。また前掲 Peterson and Augustine (2000) は、同調査の結果を分析したものである。
- 8 本調査は、21世紀COEプログラム「21世紀型高等教育システム構築と質的保証」プロジェクトの一環である「大学の組織改革についての調査」と同時に実施された。
- 9 質問紙においては、標準化型と成果準拠型に対応する根拠情報を、複数の項目に分けて尋ねた。前者には「学外の専門家による見解」、「評価機関による評価結果」、および「他大学・他学部の現状」が、後者には「卒業生の進路」、GPAなどの「学内指標における成績」、国家試験の合否などの「学外指標における成績」、および「教育成果に関する意見」が含まれる。上記ふたつの評価方法に関しては、対応する根拠情報をいずれかひとつでも利用している場合に、当該の評価方法を採用しているものとした。
- 10 文系学部には人文系・社会系・教員養成系およびそれらの組み合わせによる学際系、理系学部には理学系・農学系・工学系・保健系およびそれらの組み合わせによる学際系が含まれる。職業の想定に関しては、本調査における「学生の進路として、具体的な職業を想定している」という問いにおいて、「当てはまる」か「ある程度当てはまる」と答えた学部は「肯定」、「あまり当てはまらない」か「当てはまらない」と答えた学部は「否定」とした。入試難度に

は「第4回ベネッセ駿台共催記述模試」(2005年10月開催)の結果による合格率60%以上の偏差値を用い(朝日新聞社出版本部「大学」編集室2006、参照)、サンプルがほぼ二分される53以下を「下位」、54以上を「上位」とした。

参考文献

- 朝日新聞社出版本部「大学」編集室(2006)『2007年版 大学ランキング』、朝日新聞社
- 大山泰宏(2001)「大学教育評価の課題と展望」『京都大学 高等教育研究』、7: 37-55
- 串本剛(2006)「大学教育におけるプログラム評価の現状と課題 - 教育成果を根拠とした形成的評価の確立を目指して - 」、『大学論集』、37: 265-276
- 串本剛・渡辺明美(2006)「評価インターン活動報告書 - 広島県尾道市立栗原北小学校における学校評価表の開発 - 」(日本評価学会「学生インターン出前サービス」事業報告書)
- 葛城浩一(2006)「第4章 大学教育の成果としての性格・態度特性の変容」、広島大学・高等教育研究開発センター編『学生からみた大学教育の質 - 授業評価からプログラム評価へ - 』(COEシリーズ18)、55-70
- 清水一彦(2004)「大学評価の体系」、山野井敦徳・清水一彦編著『大学評価の展開』、東信堂、27-52
- 関正夫(1989)「教育評価の原理と方法に関する一考察」、『大学論集』、20: 3-31
- セイモア(館昭・森利枝訳2000)『大学個性化の戦略 高等教育のTQM』、玉川大学出版部
- 大学基準協会(2000)「大学評価の新たな地平を切り拓く(提言)」
- ドラッガー・スターン編著(田中弥生監訳2000)『非営利組織の成果重視マネジメント』、ダイヤモンド社
- 日本私立大学連盟 学生部会(2000)『ユニバーサル化時代の私立大学 そのクライアントの期待と要望』、開成出版
- 広島大学・大学教育研究センター編(1999)『大学改革と市場原理 第27回(1998年度)研究員集会の記録』(高等教育研究叢書 56)
- 牟田博光(2001)「革新と透明性向上のために - シンポジウム: 各界における評価の今日的課題と展望」、『日本評価研究』、1(1): 11-24
- 文部科学省高等教育局大学振興課(2006)「大学における教育内容等の改革状況について」(http://www.mext.go.jp/b_menu/houdou/18/06/06060504/001.pdf、2006/07/01アクセス)
- 米澤彰純・村澤昌崇・作田良三(2000)「日本の大学評価システムの構造と機能 - 自己点検・評価が生み出したもの」、『高等教育研究』、3: 173-192
- リースマン(喜多村和之ほか訳1986)『リースマン 高等教育論』、玉川大学出版部
- Barr, R. and Tagg J. (1995). From Teaching to Learning: a new paradigm for undergraduate education. *Change*, November/December, 13-25.
- Conrad, C. and Wilson, R. (1985). *Academic Program Reviews: institutional approaches, expectations, and controversies*. ASHE-ERIC Higher Education Reports, 5. Washington, DC: Association for the Study of Higher Education.
- Davies, S. (2002). Marketing in Higher Education: matching promises and reality to expectations. In OECD *Responding to Student Expectations*. 103-114.
- Eisner, E. (2003). Educational Connoisseurship and Educational Criticism: an arts-based approach to educational evaluation. In T. Kellaghan, and D. Stufflebeam (Eds.), *International Handbook of Educational Evaluation*. Dordrecht: Kluwer Academic Publisher, 153-166.
- Fitzpatrick, J., Sanders, J., and Worthen, B. (2004). *Program Evaluation: alternative approaches and practical guidelines*. Boston: Allyn and Bacon.
- Gardner, D. (1977). Five Evaluation Frameworks: implications for decision making in higher education. *Journal of Higher Education*, 48 (5), 571-593.
- Garvin, D. (1988). *Managing Quality*. New York: The Free Press.
- House, E. (1983). Assumptions Underlining Evaluation Models. In G. F. Madaus, M. Scriven, and D. L. Stufflebeam (Eds.), *Evaluation Models: viewpoints on educational and human services evaluation*. Boston: Kluwer-Nijhoff, 45-64.
- Huba, M. and Freed, J. (2000). *Learner-Centered Assessment on College Campuses: shifting the focus from teaching to learning*. Boston: Allyn and Bacon.
- Palomba, C. and Banta, T. (Eds.) (1999). *Assessment*

- Essentials*. San Francisco, Calif.: Jossey-Bass.
- Peterson, M. and Augustine, C. (2000). Organizational Practices Enhancing the Influence of Student Assessment Information in Academic Decisions. *Research in Higher Education*, 41 (1), 21-52.
- Peterson, M., Einarson, M., Augustine, C., and Vaughan, D. (1999). *Institutional Support for Student Assessment: methodology and results of a national survey*. Stanford, Calif.: National Center for Postsecondary Improvement.
- Rogers, P., Hacsí, T., Petrosino, A., and Huebner, T. (Eds.) (2000). *Program Theory in Evaluation: challenges and opportunities*. (New Directions for Evaluation, 87) San Francisco, Calif.: Jossey-Bass.
- Stark, J. and Lattuca L. (1997). *Shaping the College Curriculum: academic plan in action*. Boston: Allyn and Bacon.
- Stufflebeam, D. (2001). *Evaluation Models*. (New Directions for Evaluation, 89) San Francisco, Calif.: Jossey-Bass.
- Trigwell, K. and Prosser, M. (1996). Changing Approaches to Teaching: a relational perspective. *Studies in Higher Education*, 21 (3), 275-284.
- Tyler, R. (1949). *Basic Principles of Curriculum and Instruction*. Chicago: The University of Chicago Press.

(2007.2.1受理)

Diversification of Approaches to the Self-Evaluation of Undergraduate Education

Takeshi Kushimoto

Hiroshima University Graduate School of Education
Japan Society for the Promotion of Science Research Fellow
t.kushimoto@hiroshima-u.ac.jp

Abstract

The purpose of this study is proposing four evaluation methods and demonstrating to what degree they are adopted in self-evaluations of undergraduate education in Japan. In the first part, four evaluation approaches are introduced in relation to the ideas of undergraduate education. They are named as Autonomy-Based Approach, Demand-Based Approach, Standard-Based Approach, and Outcomes-Based Approach. In the latter part of this article, the data from a national survey of 1,871 faculty members in Japanese universities show that to what extent those four methods are used in practice. As a result of the analysis, it is found that the degree of adoption depends on the evaluation focuses and traits of the faculty. This fact is discussed in consideration of diversifying idea of undergraduate education.

Keywords

undergraduate education, self-evaluation, evaluation methods,
diversification, national survey