

研究成果に対する主観的評価の構造と測定可能性

加藤 毅 (筑波大学)

1. はじめに

自己点検・評価の義務化（平成11年）に続いて、平成15年4月からは認証評価機関による評価が大学に対して義務づけられた。これと並行して、大学評価・学位授与機構のもとですでに3回に渡って評価結果が確定、公表されている。また、平成14年度には「21世紀COEプログラム」が新設され、平成15年度には56大学133の研究教育拠点が新規に採択された。これらの制度に対する指摘の多くは、評価にかかわる問題点や今後の課題など批判的観点からのものである(注1)。

研究活動の評価において用いられる中心的な手法がピア・レビューである(注2)。大学において行われる活動のなかで、個人（研究グループ）による研究活動については、大学審議会が設立されるはるか以前から広く評価の対象とされてきた。これまでの経験とそこでの蓄積を活かして『文部科学省における研究及び開発に関する評価指針』（平成14年6月）が設定され、現在あるべき評価手法について模索が続けられている。大学における学術研究を評価することの難しさについては、同指針のなかでも強調されているように、すでに広く認知されているとあってよい。その一方で、例えば論文数など入手可能な定量的データのみに基づいて研究活動を評価することも可能である(注3)。そこでの分析結果に対して、これまで「本来の成果のうち測定しやすい部分を強く反映した偏ったもの」（塚原 2003）であるとの指摘が繰り返されてきた(注4)。大学の研究評価もまた、「評価の難しさと易しさが共存」（塚原 2003）する状態にあるとあってよい。

21世紀に入り大学改革が「競争と評価」の局面（天野 2002）を迎えつつある。比較的容易に行うことができるが故に、本格的なピア・レビュー評価に先立って定量的データに基づく研究評価を行おうとする場合、これまででない新たな問題に直面することになる。

第一の問題点は、定量的な評価はこれまで専ら自然科学分野に対して行われてきており、現在用いられている手法では、研究成果の発表スタイルが大きく異なる人文社会分野の測定が困難(注5)であることによるものである。人文社会分野もいまや評価の圧力から自由ではあり得ず、自然科学分野を前提とした手法の安易な適用を避けるためには、みずからの独自性について積極的に提言していく必要がある。そのためにもまず、研究成果の発表スタイルの独自性について定量的に明らかにする取り組みが必要である。

第二に、例えば独創性ゆえに支配的な基準に適合しにくいケースに代表されるように、自然科学分野のなかにあっても評価基準は必ずしも合意されているわけではない(注6)。定量的な評価が「第三世代の大学評価」（金子 2000）のもとで資源配分に大きく影響するようになった場合の副次効果は計り知れない。この分野においても、定量的な評価の問題点を指摘すること

どまらず、具体的な代替案の策定に向けた取り組みが必要ではないか。

第三に、政府や第三者機関によるこのような評価活動そのものが大学の活動全体に及ぼす影響について取り上げたい。およそ「評価や審査の公平性、ひいては評価者にたいする根強い不信の念」（天野 2002）が隠されたなかで行われる評価を通じて、「大学等の教育研究水準の向上に資する」（「大学評価・学位授与機構」）あるいは「世界最高水準の大学づくりを推進する」（「21世紀COEプログラム」）といった目的が達成されるとは考えにくい。不信を増幅するのではなく「研究評価の結果が、研究の発展に結びつきその推進力となること」（中井 1995）こそが重要なのであり、そのためには「可能な限り、多くの研究者が評価作業に参加する機会を持つこと」「広く研究者集団の意見を聴き、集約して評価に活かすこと」（中井 1995）が必要なのではないか。「研究者の意見が集約され、それが具体的に反映されること」が、健全で効果的な評価システムを築く道」（中井 1996）であるとの提言は、研究者集団に信頼されそして受容される評価基準の必要性を説くものにほかならない。

このような問題関心から本稿では、わが国の大学教員によって内面化されている研究成果の発表形態に対する主観的な評価について、実証的分析を試みる。次節ではまず、研究成果の発表形態や評価基準が専門分野によってどのように異なるのかを定量的に明らかにする。それぞれの分野の平均的な姿をふまえたうえで、第3節以降では、主成分分析を用いてそれぞれの分野の評価構造の特徴を明らかにしたうえでそれらの共通点についても探る。同時に、主観的な評価が研究成果の発表形態選択に及ぼす影響についても検討を行う。最後に、分析の結果をふまえ、今後の研究評価のあり方についての政策的検討を行う。なお、分析に用いた調査の概要は表1に示す通りである（注7）。

表1 調査の概要

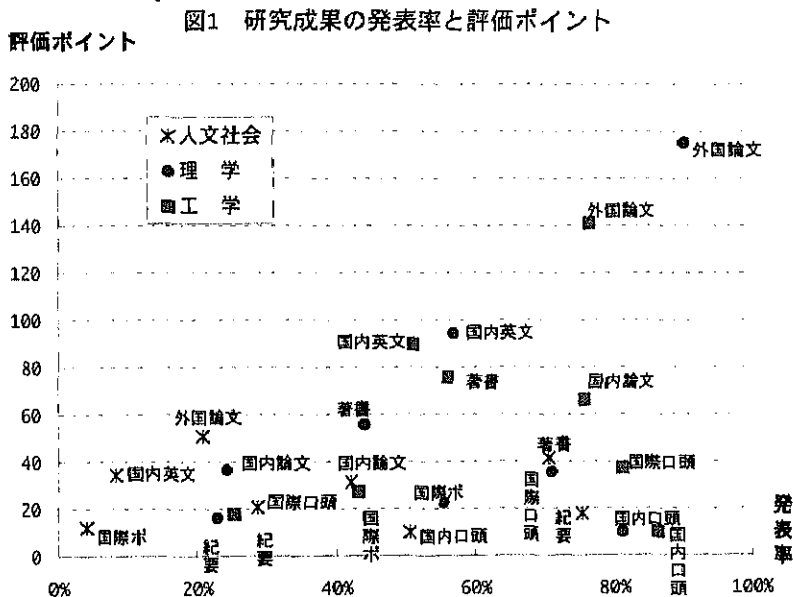
調査対象	日本全国の国・私立四年制大学の人文、社会、理、工、農、薬およびその他の学部・大学院、附置研究施設等の専任の教授および助教授を対象とし、10,600サンプルを無作為に抽出（抽出率は23%）
調査時期	1998年11月下旬～1999年1月
調査方法	郵送による配付・回収
回収率	29%（3,075票）

2. すれ違う評価基準

2-1. 異なる「外国論文」の位置づけ

専門分野によって研究成果の発表形態はどのように異なっているのか。また、それぞれの発表形態に対する主観的な評価はどの程度異なっているのか。ここでは代表的な研究成果の発表形態として、外国の審査つき学術論文（外国論文）、国内の審査つき英文学術論文（国内英文）、国内の審査つき和文学術論文（国内論文）、国際学会での口頭発表（国際口頭）、国内

の学会での口頭発表（国内口頭）、国際学会でのポスター発表（国際ポ）、大学の紀要等に発表した研究論文（紀要）、著書を取りあげ（注8）、人文社会、理学、工学からなる3つの専門分野について比較を行った（図1）。図の横軸は、当該形態において過去3年間に1編以上発表を行ったものの比率（発表率）であり、研究成果を発表する場としての認知度や普及度を示すものと考えられる。縦軸は、「国内口頭」を10ポイントとしたときのそれぞれの相対的な評価（評価ポイント）を示している。図の右上に位置するのは発表率・評価ポイントとも高い発表形態であり、質・量ともに研究活動の中心となっていると考えられる。図の左下は発表率・評価ポイントともに低く、発表形態としてはあまり機能していないことを示す。右下にくるのは発表率が高く評価ポイントが低い発表形態であり、これは発表の機会が広く開かれているものと考えられる。



まず理工系分野についてみると、もっとも高い評価ポイントが与えられているのが「外国論文」であり、理学は175ポイント、工学でも141ポイントと圧倒的である。発表の場としての普及度も高く、理学では90%、工学でも76%の回答者が過去3年間に「外国論文」の発表を行っている。他方「国際口頭」や「国内口頭」は、発表の場として選ばれることが多くなっている一方で、相対的な評価は低い（40ポイント以下）。これらのことから理学や工学では、優れた研究成果は「外国論文」、その他の成果については国内外の学会での口頭発表という形態が主要な発表の場として活用されていることがわかる。

人文社会の最大の特徴は、最も評価の高い「著書」ですらその評価ポイントは基準とした「国内口頭」の4倍にすぎないという点である。発表形態間の評価ポイントの差は一般的に小さく、理工系で圧倒的に高い評価を受けている「外国論文」の評価ポイントは51、発表率もわずか21%に過ぎないマイナーな存在となっている。学会での口頭発表率も同様に低く、これに代わって理工系分野ではほとんどみられない「紀要」が主要な発表形態となっている。

2-2.人文社会における外国論文

ここで、人文社会における「外国論文」の評価と発表率について考えてみよう。51ポイントという数字はあくまで回答者全員の平均であり、なかには「外国論文」に対して非常に高いポイントを与えるものも存在する。「国内論文」と「外国論文」の評価ポイントの大小関係に着目し、「外国論文」に対する相対的評価が高いものの「外国論文」発表率が高く、同時に、「国内論文」の評価が高いものでは「外国論文」の発表率が低くそして「国内論文」の発表率が高くなっていったとしよう。この場合には主観的評価の高い形態が成果発表の場として選択されていると考えることができる。他方もし「外国論文」に対する相対的評価が高いにもかかわらず「外国論文」の発表率が低いままであるとすれば、主観的評価が発表形態の選択に影響を及ぼしているとは考えにくい。

表2によれば、「外国論文」よりも「国内論文」を高く評価するものでは、「国内論文」に発表していて「外国論文」には発表していないものの比率が41%と高く、「外国論文」のみはみられない。これに対して、「外国論文」に対する相対的評価が高くなるにつれて「外国論文」の発表率は増加し、その差が100ポイントを越えると「外国論文」の発表率は40%に達する。そのうち「国内論文」を発表したものに絞ってみると、その半数は「外国論文」を発表していることになる。この結果は、主観的評価の高い発表形態が積極的に選択されているという仮説を支持するものである。

表2 外国・国内論文の相対的評価ポイントと成果発表（人文社会）

<評価ポイント>	外国論文を発表		外国論文なし
	国内論文なし	国内論文を発表	
外国論文<国内論文	0%	18%	41%
外国論文=国内論文	6%	9%	36%
外国論文>国内論文	10%	12%	29%
外国論文>100p.	15%	25%	24%

このような状況のもと、「外国論文」に対して圧倒的に高い評価を与える理工系の基準に基づいて人文社会の研究者の評価が行われたとしよう。そこでは、結果的に「外国論文」を高く評価しそして発表形態として選択している研究者が機械的に高く評価されることになる。つまり、研究成果の質や研究者の力量とは無関係に、評価対象者の主観的評価基準そのものが評価対象となるわけである。このような評価基準が人文社会の研究者によって受け入れられるとは考えにくい。

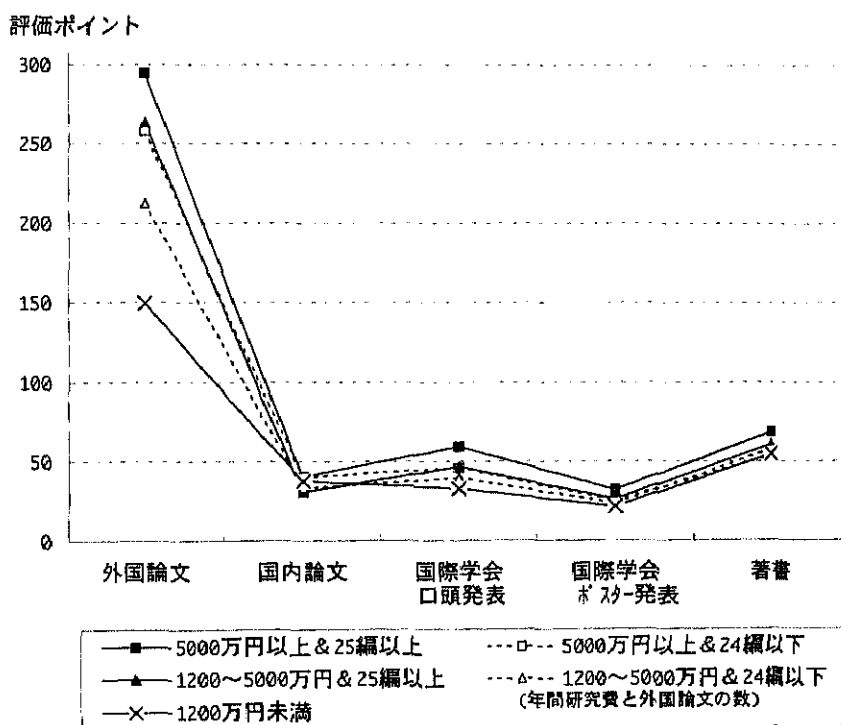
2-3.評価主体の重点・肯定的な評価観

専門分野の間にみられる主観的評価基準の相違はすでに示した通りであるが、分野を同じくするからといって必ずしも評価基準が共有されているわけではない。ここでは、研究成果に対して評価を行う立場にある研究者と評価対象とされる研究者一般の間で評価基準は共有されて

いないのではないか、という仮説を設定し、理学分野を対象として分析を行った(注9)。なお、研究評価を行う立場にあるか否かを直接的に判断することは難しいため、ここでは、研究費の規模および研究成果(「外国論文」)の多寡をその代理指標として用いた。

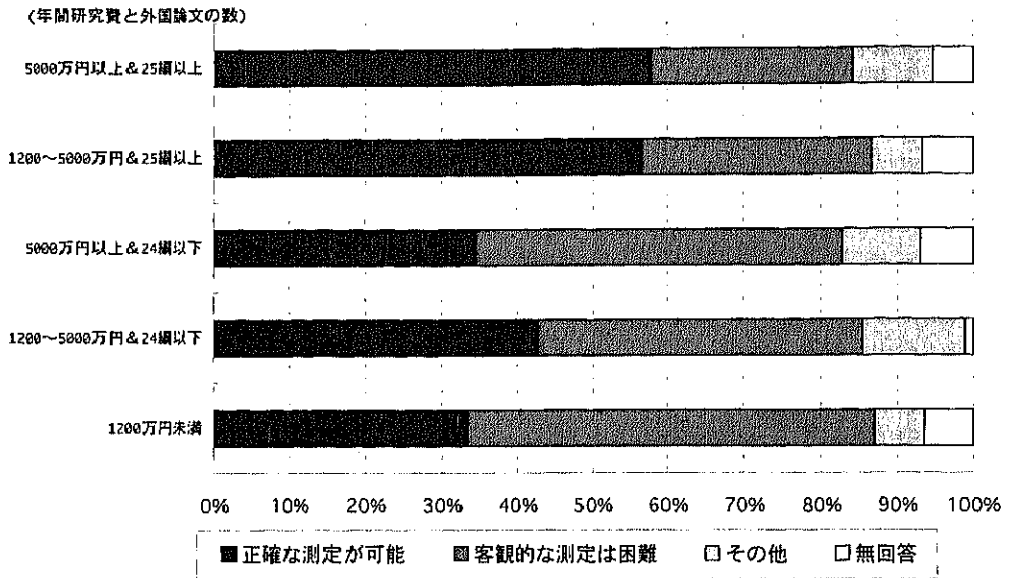
年間研究費と外国論文数の組み合わせ別に、それぞれの発表形態の評価ポイントをみたものが図2である。年間研究費が最も多く(年間5,000万円以上)、同時に多数の優れた論文を發表している(過去3年間に「外国論文」が25編以上)ものは、「外国論文」に対して295ポイントという非常に高い主観的評価を与えている。これに対して、全体の66%を占める年間研究費が1,200万円未満のものあいだでは、「外国論文」に対する主観的評価はかなり低く150ポイントでしかない。つまり、同じ専門分野のなかにあっても評価基準が共有されているわけではなく、評価を行う立場にあると考えられる研究者は、一般の研究者と比べて「外国論文」を選択的に非常に高く評価していると考えられる。

図2 立場によって異なる評価ポイント(理学)



さらに重要であると考えられるのが、研究費や研究成果の多寡によって研究評価に対する考え方がかなり異なるという点である。研究成果の客観的な測定可能性に関する質問に対する回答をみると(図3)(注10)、5,000万円以上で25編以上、あるいは1,200万円~5,000万円で25編以上のおよそ60%が「正確な測定が可能」と答えており、「客観的な測定は困難」は30%に満たない。これに対して年間研究費1,200万円以下(全体の66%)では両者の比率は逆転し、「客観的な測定は困難」が過半数を超えている(54%)。

図3 立場によって異なる評価観（理学）



理学分野ではこのように、評価を行う立場にあると考えられる研究者は、「外国論文」という発表形態を選択的に高く評価すると同時に、その定量的な評価を通じて研究成果を正確に測定できると考える傾向が強い。この主観的評価基準および評価観は、評価対象とされる研究者一般のものからはかなり乖離している（評価基準や評価観は共有されていない）のである。

3. 人文社会の四類型 ～主観的評価基準と発表形態の選択～

ここまで、専門分野や立場の違いによる評価基準のズレとその問題点に関する議論のなかでは、全体的な傾向を明らかにするという目的から主として平均値に着目した分析を行ってきた。もちろんデータを細かくみていくと、同じ専門分野のなかでもかなりのばらつきが観察される。本節では、研究成果の主観的評価の全体像について構造的に把握することを目的として、専門分野別に主成分分析を行なう（注11）。

まず人文社会を対象とする分析の結果（表3）をみると、第一主成分の固有ベクトルがすべて正值となっている。このことから、析出された第一主成分は主として「外国論文」「国内論文」「著書」などに対する評価の高さを表していることがわかる。なお、分析に用いた評価ポイントは「国内口頭」を基準とする相対的なものであることから、第一主成分得点が低くなるのは、発表形態による評価ポイントの差が小さい場合である。この主成分の寄与率は高く64%に達する。

人文社会の第二主成分をみると、「外国論文」や「国際口頭」に対する固有ベクトルが負の値を示し、「紀要」「著書」「翻訳」が正值となっている。この主成分は、発表に用いた言語に対応しており、第二主成分得点が高い場合には日本語での発表が、得点が低い場合には外国語での発表に対して高い評価が与えられていることを示す。第二主成分までの累積寄与率はお

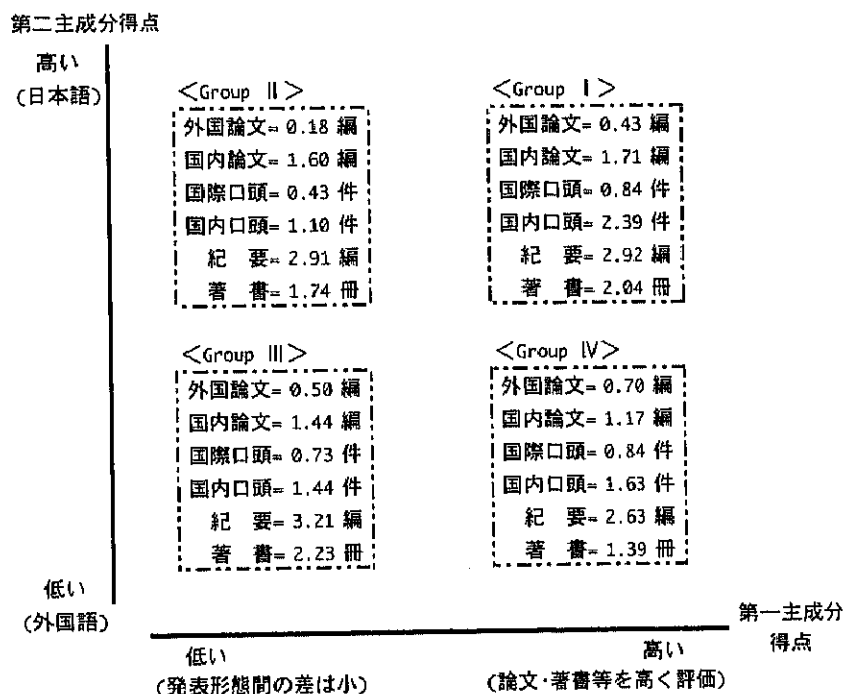
表3 研究成果の主観的評価の全体構造

		＜固有ベクトル＞									
		固有値	累積 寄与率	外国論文	英文論文	国内論文	国際学会 口頭発表	国際学会 ポスター発表	紀要	著書	翻訳
＜人文社会＞											
第一主成分	3.869	64%	0.45	-	0.46	0.36	-	0.36	0.43	0.39	
第二主成分	0.984	80%	-0.35	-	-0.08	-0.65	-	0.49	0.23	0.39	
＜理学＞											
第一主成分	3.407	57%	0.46	0.48	0.39	0.39	0.35	-	0.36	-	
第二主成分	1.039	74%	-0.09	-0.21	-0.39	0.54	0.61	-	-0.36	-	
第三主成分	0.644	85%	-0.52	-0.38	0.21	0.00	0.31	-	0.66	-	
＜工学＞											
第一主成分	4.073	68%	0.45	0.45	0.42	0.40	0.38	-	0.34	-	
第二主成分	0.799	81%	-0.22	-0.33	-0.33	0.54	0.63	-	-0.21	-	

よそ80%に達する。人文社会では客観的な評価の可能性に対して否定的な見解が多いことについてはすでに触れたが、その一方で、定量的データに対する主観的評価基準そのものについては比較的単純な構造を持っていることがわかる（注12）。

それでは、このようにパターン化された主観的な評価基準は、成果を発表する際の行動（発表形態の選択）にどのような影響を及ぼしているのだろうか。評価パターンと成果の発表形態との関連について分析したものが図4である。図の右下に位置するのは、第一主成分得点が高く第二主成分得点の低い、すなわち「国内口頭」以外の発表形態が相対的に高く評価され、同

図4 評価パターンと研究成果（人文社会）



時に外国語での発表を高く評価するグループである（グループⅣ）。このグループの過去3年間の研究成果は自らの主観的評価に対応しており、高く評価する「外国論文」が0.70編、「国際口頭」が0.84件と多く、評価の低い日本語での発表（「国内論文」「紀要」「著書」）は少ない。

図の左下に位置するのは、両主成分得点ともに低い、すなわち発表形態間の差が小さく、英語での発表を評価するグループ（グループⅢ）である。このグループの研究成果は、グループⅣに比べて、「外国論文」「国際口頭」がやや少なく、「国内論文」や「紀要」「著書」などの発表件数がやや多くなっている。この傾向をさらに強めたものが図の左上に位置するグループⅡであり、「外国論文」「国際口頭」の発表件数が少なく「国内論文」が多くなる。そして、図の右上に位置するグループⅠでは、「国内論文」「国内口頭」を通じた発表がもっとも多く行われている。

人文社会ではこのように、主成分分析を通じて析出された主観的評価のパターンに対応して研究成果の発表が行われているといえる。この評価パターンはさらに、研究成果の評価に対する考え方に対しても影響を及ぼしている。「正確な測定が可能」と答えるものの比率がもっとも高いのは、外国語を重視したメリハリのある評価基準を有するグループⅣ（15%）である。対照的な評価基準を有するグループⅡでは、その比率はもっとも低く8%未満となる。

4. 相違のなかの安定的な評価構造

すでにみたように、人文社会と理工系分野では、主観的評価における「外国論文」の重みが大きく異なる（図1）。すなわち、人文社会では発表形態間の評価ポイントの差はそれほど大きくなく、そのうえ「国内論文」を「外国論文」以上に高く評価するものの比率も27%に達する。これに対して理工系分野では、「外国論文」に対して圧倒的に高い評価ポイントが与えられており、たとえば理学における「外国論文」の評価ポイントの平均値は「国内口頭」の17倍を越えており、また「外国論文」に対して「国内論文」の5倍以上の評価ポイントを与えるものも全体の30%に達する。

その一方で主成分分析を通じて、異なってみえるこれらの専門分野の間にかなり共通した評価構造が存在することが明らかになった。理学および工学における主観的評価の構造（表3）についてみると、両分野に共通して第一主成分の固有ベクトルがすべて正值となっており、その寄与率もかなり高い（理学で57%、工学では68%）。このように、「国内口頭」に対する「外国論文」や「国内論文」等の相対的評価の高さを強く反映した第一主成分は、専門分野を問わず安定していることがわかる。しかもその寄与率は大きく、第二主成分以下を圧倒している。

第二主成分以下についても比較検討を行おう。理学における第二主成分は「国際口頭」や「国際ポ」に対する固有ベクトルが正值、その他の国内外での論文や著書は負値をとっていることから、国際学会での発表に対する評価の高さを反映したものであると考えられる（第二主成分までの累積寄与率は74%）（注13）。さらに第三主成分をみると、「外国論文」と「国内

英文」が負値、その他は正值であることから、英文（外国語）の論文に対する評価を反映したものと考えられる。最後に工学についてみると、第二主成分は理学分野と同様に国際学会での発表に対する評価をしめしており、累積寄与率は理学よりもやや高く82%となっている。

さらに、「外国論文」に対して高い評価を与えるものほど研究成果の「正確な測定が可能」と応えるものの比率が高くなっている。すなわち、理学全体では38%が「正確な測定が可能」と回答しているのに対して、「外国論文」の評価が50ポイント未満では33%、200～500ポイントでは40%、そして500ポイント以上となるとその比率は53%に達している。同様の傾向は、工学についても観察される。

5. 工学における国内論文の位置づけ

理学と工学の間にみられる大きな相違の一つが「国内論文」の位置づけである。「国内論文」に対する評価は工学でもっとも高く66.4ポイント、同時に過去3年間に「国内論文」に発表したことのあるものの比率（発表率）も76%と高い。理学においてはマイナーで評価の低い発表形態となっているのとは対照的に、工学では「国内論文」は重要な成果発表媒体として機能しているのである。この点についてさらに考察を進め、「国内論文」に対する「外国論文」の相対的評価ポイントに着目して工学における研究評価のあり方について検討を行った。

表4は、「国内論文」と「外国論文」の評価ポイントの大小関係と研究成果の関連についてみたものである。「外国論文」を6編以上発表しているものの比率をみると、「国内論文」が「外国論文」以上に評価されているグループでは7%、「外国論文」のポイントが「国内論文」の2倍未満では11%、5倍未満では22%、そして5倍以上のグループではその比率はさらに高く42%となっている。「外国論文」を発表していないものの比率をみると、「外国論文」のポイントが「国内論文」の2倍未満では25%を越えているのに対して、2倍以上では21%、そして5倍以上ではわずか5%に過ぎない。このように、「国内論文」に対する「外国論文」の相対的な評価ポイントが高くなるほど「外国論文」の発表数は多くなっているのである。他方「国内論文」の発表数についてみると、「外国論文」とは対照的に、「外国論文」の相対的な評価ポイントが高いほど発表数は少なくなっている。さらに、「外国論文」を相対的に高く評価するものほど研究成果の「正確な測定が可能」とするものの比率は高く、「外国論文」の評価ポイントが「国内論文」の5倍以上のグループでは45%に達する。

表4 外国・国内論文の相対的評価と論文数（工学）

<評価ポイント>	外国論文の発表数（過去3年間）					国内論文の発表数（過去3年間）					
	6編～	4-5編	2-3編	1編	0編	11編～	7-10編	5-6編	3-4編	1-2編	0編
外国論文≤国内論文	7%	9%	19%	37%	27%	12%	17%	19%	22%	20%	9%
外国論文>国内論文	11%	12%	20%	30%	28%	15%	10%	14%	17%	22%	21%
外国論文≥国内論文×2	22%	18%	14%	24%	21%	8%	12%	13%	15%	23%	29%
外国論文≥国内論文×5	42%	21%	14%	17%	5%	1%	3%	9%	8%	14%	64%
全体	16%	14%	17%	29%	24%	11%	12%	15%	17%	21%	24%

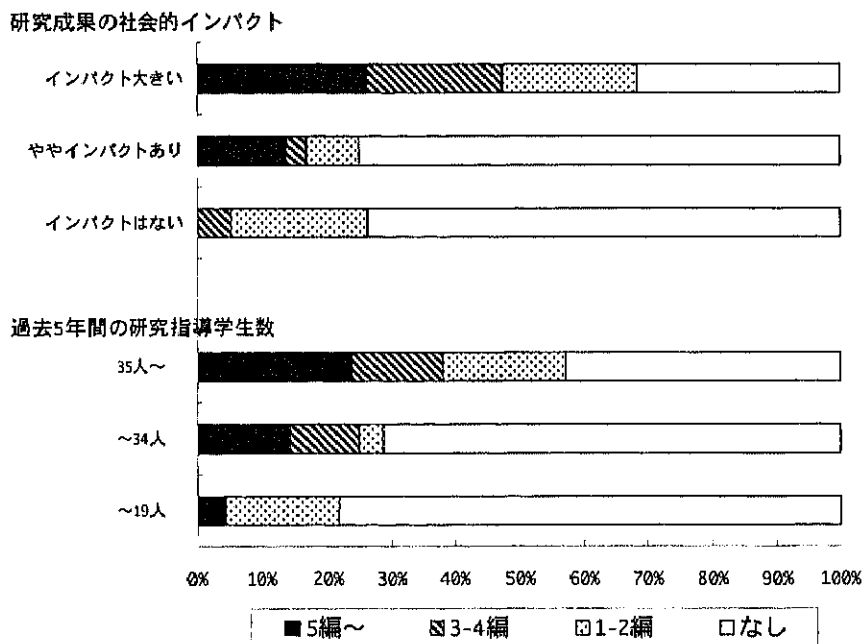
つまり、「外国論文」を高く評価するものは、自身の評価基準にもとづいて「外国論文」という形態での成果発表を積極的に行っているとともに、「外国論文」を重視した研究評価に対してもかなり肯定的である。もしもここで「外国論文」に大きなウエイトをおく客観的な評価が実施された場合、「国内論文」を相対的に高く評価するものは、主観的な評価基準の違いがそのまま反映された結果として低い評価を受けることになる。

6. 多目的化する研究活動と評価基準

それでは、「外国論文」を高く評価するものにとって「国内論文」はどのような意味を持っているのであろうか。「外国論文」に対して「国内論文」の5倍以上という高い評価ポイントを与えているグループのなかの36%は「国内論文」を発表している（表4）。評価の低さが発表率の低さに連動している理学とは異なり、工学では、相対的評価が低くなっても依然として「国内論文」を通じた成果の発表が行われている。ここで考えられるのが、「国内論文」が、学術研究の成果を競う場としてではなくそれ以外の大学の多様な機能を発揮する場として活用されているのではないか、という仮説である。例えば、専門的な学会ではなく産業界やさらには社会に対して研究成果を発表する場として、あるいは大学院生に対する研究指導の一環として「国内論文」という成果発表の場を活用することは、十分に可能である。

この仮説について検討を行った結果が図5である。ここでは、「外国論文」の評価ポイントが「国内論文」の5倍以上のグループをとりあげ、研究成果の社会的インパクトの大きさ及び過去5年間に研究指導を行った学生（大学院生を含む）の人数と「国内論文」数との関係につ

図5 社会貢献・教育活動と「国内論文」数（工学）



いて分析を行った。たとえば「大きなインパクトを与えた」とこたえるものでは26%が5編、21%が3-4編の「国内論文」を発表しているのに対して、「ほとんど与えなかった」ものでは5編以上が皆無、3-4編はわずか5%であり、74%は「国内論文」を発表していないのである。学生への研究指導との関係のみでも、指導学生数が多くなるほど「国内論文」の発表編数が多くなっていることがわかる。この結果は、「国内論文」という成果発表媒体が社会貢献活動や教育活動の一環として活用されているという仮説を、直接的ではないにしても支持するものである。

大学に対して、従来の伝統的な研究後継者養成にとどまらず知識に関わるあらゆる場面で教育的機能を果たすことが期待されており、また「科学のモード論」（ギボンズ他, 1994）では、研究活動と社会貢献活動の融合をモチーフとする議論が展開されている。この議論の延長上に、研究評価のあり方をめぐる新たな論点が浮かび上がる。もし研究成果の発表が、専門分野の研究の発展に直接貢献するという限定的な位置づけを越えて、教育あるいは社会貢献活動の一環として重要な機能を果たしているとしよう。このとき研究活動は、狭い意味での研究にとどまらず大学の活動全体に対する評価のなかに位置づけられることが必要となる。つまり研究活動が教育や社会貢献活動と融合したとき、研究評価は、純粋な学術の評価にとどまらず大学に期待される他の諸機能を含む広い文脈のなかで行われることが必要となるのではないか(注14)。

7. まとめと政策的インプリケーション

本分析を通じて、研究成果の主観的評価基準や客観的な評価の可能性について、分野や立場によってどのように異なるのか、その実態が定量的に明らかになった。評価基準の大きな相違は分野を越えた定量的評価が著しく困難であることを示し、さらに分野によっては主観的評価が成果の発表形態に結びついているケースも存在する。また、評価を行う立場にあるものとその他の研究者一般の間にみられる評価基準や評価観のギャップは、これまでの評価のあり方を根本から問うものである。判断の基準が受容されないままに資源配分にリンクした評価が実施されるようなことがあれば、不信感はさらに増幅され、我が国の学術研究システムは大きく歪められることにもなりかねない。

分析の結果は、研究成果を定量的に評価することの困難性を示すと同時に、新たな可能性や今後の方向性についても示唆を与えてくれる。まず第一に、研究成果に対する評価の基本的構造には、分野を越えた類似性が存在することが明らかになった。これらの類似性は、複数の分野にまたがって行われる評価の基礎をなすものである。意見の集約をめざして、異なる評価基準を有する研究者による建設的な議論の積み重ねが求められよう。とくに理学や工学では、客観的な研究成果の測定に対する肯定的意見も多く、研究者集団に受容されそして信頼されるような定量的データを用いた評価基準の構築は、「易しい」ことではないけれども、決して不可能ではない。

本稿が分析対象としてきた研究成果の定量的評価は、いままでもなく研究評価の一部分を構

成するものでしかない。しかも、大学改革がめざす「知識基盤社会における高等教育」(文部科学省 2004)においては、研究活動に対する評価は従来のような学術的な観点に限定されるものではあり得ない。しかしながら、公平で透明な評価基準を作成する上での土台となりうるものであるからこそ、定量的な評価基準を巡る今後の可能性に対する期待は高い。健全で効果的な評価システムの構築に向けて、残された課題は多い。

<注>

- 注1.「大学評価の新段階」をテーマとする『IDE現代の高等教育』No.442(2002)のなかで、様々な立場から、大学評価の在り方について多くの問題点が指摘されている。また、対象機関および関係団体から、多数の意見が大学評価・学位授与機構に対して寄せられている(大学評価・学位授与機構 2002, 2003)。評価の基準や手法を超えて大学評価そのものが抱える問題やその背景等については、金子(2000)や天野(2002)などを参照。21世紀COEプログラムについては、たとえば読売新聞大阪本社(2003)や岩崎ほか(2002)など。なお、資源配分に際して用いられる研究の評価に対する批判的見解は、最近にはじまったことではなく以前からもしばしば指摘されてきたものである(松尾 1995、野依 1996、加藤 2000)。
- 注2.ピア・レビューとは、「同一あるいは近接する専門分野の科学者によっておこなわれる科学的な価値についての判断」をさす(OECD 1987)。その詳細については、たとえばKostoff(1997)や山崎(2001b)など。
- 注3.この手法を用いた我が国の大学の分析結果については、たとえば慶伊ほか(2000)や孫(2001)など。
- 注4.たとえば、根岸ほか(2000)や慶伊ほか(2000)、山崎(2001a)など。研究評価の先進国とも言える米国における近年の公的研究開発の評価をとりあげた研究のなかでも「万能な評価手法はない」ことが指摘されている(富澤ほか 2002)。
- 注5.問題の所在自体については、これまで太田和(1996)、大場(2000)、根岸ほか(2000)などを通じて指摘されているが、ここで重要なことはその定量的実態を明らかにすることである。
- 注6.レフェリー・システムの保守性についてはすでに指摘されており(山崎 2001b)、定量的データをもとに研究評価を行う場合にも同じような問題が生じることが危惧される(大学基準協会 2000)。
- 注7.調査の概要と全般的な分析結果については、『大学研究』第21号を参照。
- 注8.人文社会を包含していることや、外国のジャーナルに発表した論文を高く評価する理工系の慣行を考慮して、ここでは計量書誌学でしばしば用いられる米国ISI社の引用統計データベースが採用する文献種別(Article, Note, Proceedings, Review, Letter, Editorial,

Others)とは異なる分類を設定した。また、分野を越えた比較はあまり意味がないなどの理由により(山崎 2001a)、インパクト・ファクターについては取り扱わない。

注9.理学分野では「外国論文」に対する評価ポイントが圧倒的に高くなっているため、その発表編数をもって研究者の実績の有無(すなわち評価を行う立場にあるか否か)について比較的判断を下しやすい。

注10.研究成果の客観的な測定の可否に関するこの質問では、1.サイテーションインデックスやインパクトファクターを用いた研究論文の評価を通じて、研究の成果をほぼ正確に測定することができる、2.研究の成果について客観的な測定を行うことは困難である、3.その他、という3つの選択肢が設定されている。

注11.主成分分析を行うにあたり、分布の特性を考慮して評価ポイントの対数変換を行った。また、研究成果を発表する場としてほとんど用いられていない形態(人文社会では国内英文と国際ポスター、理学と工学では紀要と翻訳)については分析対象外とした。

注12.専門分野の小分類別に比較を行ったところ、法政治学ではやや論文や著書等の相対的評価が低く(第一主成分の値が小さい)、人文科学では逆に論文や著書等の相対的評価がやや高くなっていた。また、経済・商学では、外国語での発表に対する評価がやや高くなっていた。ただしこれらの差異は、ここまでの分析結果に影響を与えるような大きなものではない。

注13.第二主成分における人文社会と理工系分野の間の相違は、専門分野間にみられる発表形態自体の違い(注12を参照)を反映したものとなっている。

注14.評価される立場にある大学や研究者集団が、説明責任を負うところの社会によって有効性を認められた評価の側面を否定することは難しい。科学研究が社会から評価されるということのなかには、大学や研究者集団の内部で行われる評価活動自体に対する評価も含まれるからである。実際に、公的資金の支援を受けた基礎研究に対する評価項目として、米国ではクオリティ、世界的なリーダーシップに加えて研究のレリバンスがあげられている(COSEPUP 1999)。

<引用・参考文献>

天野郁夫 2002,「秩序の崩壊と創造」『教育学年報9 大学改革』世織書房, 5-48頁。

COSEPUP 1999, "Evaluating Federal Research Programs", NAP.

大学基準協会 2000,『大学評価の新たな地平を切り拓く(提言)』35-40頁。

大学評価・学位授与機構 2002,「平成12年度着手の大学評価に関する意見とその対応について」。

大学評価・学位授与機構 2003,「平成13年度着手の大学評価に関する意見について」。

岩崎稔,鈴木公,田栗正章 2002,「COE狂想曲」『世界』第708号, 222-232頁

加藤毅 2000,「認知された資金供給システム」『大学研究』第21号, 135-144頁。

金子元久 2000,「大学評価のポリティカルエコノミー」『高等教育研究』第3集, 21-41頁。

- 慶伊富長,本多卓也 2000,「自然科学系の研究」『高等教育研究』第3集,63-79頁.
- Kostoff R.N. 1997, "THE HANDBOOK OF RESEARCH IMPACT ASSESSMENT, 7TH EDITION" Office of Naval Research.
- 松尾義之 1995,「文部省科学研究費のゆくえ」『日経サイエンス』第284号,18-30頁.
- Michael Gibbons 1994, "The New Production of Knowledge", Sage publications.
- 文部科学省 2004,『平成15年度文部科学白書』,国立印刷局,5-7頁.
- 中井浩二 1995,『我が国における研究評価手法の総合的研究』平成5~6年度科学研究費補助金(総合研究(A))研究成果報告書
- 根岸正光 2001,「研究評価とビブリオメトリックス」根岸・山崎編著『研究評価』丸善,1-17頁.
- 根岸正光,孫媛,山下泰弘,西澤正己,柿沼澄男 2000,「我が国の大学の論文数と引用数」『学術月報』53(3),64-80頁.
- 野依良治 1996,『学術研究の実体に関する調査研究』平成6~7年度科学研究費補助金(総合研究(A))研究成果報告書.
- OECD 1987, "Evaluation of Research", pp.26-28.
- 大場淳 2000,「人文社会科学の研究環境とその評価」『高等教育研究』第3集,81-105頁.
- 太田和良幸 1996,「『学術研究の評価』に関する有識者の意見」『学術月報』Vol.49, No.2, 65-70頁.
- 孫媛 2001,「わが国の大学の論文生産と引用数」根岸・山崎編著『研究評価』丸善,69-111頁.
- 富澤弘之,斎藤芳子,小林信一 2002,『米国における公的研究開発の評価手法』文部科学省科学技術政策研究所.
- 塚原修一 2003,「教育政策と評価のダイナミズム」『教育社会学研究』第72集,5-20頁.
- 山崎茂明 2001a,「指標としてのインパクト・ファクター」根岸・山崎編著『研究評価』丸善,55-68頁.
- 山崎茂明 2001b,「研究評価のためのレフェリー・システム」根岸・山崎編著『研究評価』丸善,149-158頁.
- 読売新聞大阪本社 2003,『大学大競争』中央公論新社.

Structure and Availability of Quantitative Evaluation of Academic Research

Takeshi KATO, University of TSUKUBA

As is symbolized by the stipulation "National Guideline on the Method of Evaluation for Government R&D" and the establishment of the National Institution for Academic Degrees as an evaluation organization for universities, the mechanism for the evaluation of academic research activities is being developed rapidly. In contrast, method and standards for evaluation are not opened yet in detail. Therefore, it is quite difficult for academics to have strategies for evaluation.

It is essential for fair evaluation to be opened the method and criterion of evaluation. And it is much more important for the "encouraging academics" evaluation to have a standards in conformity with the internalized standards of academics to be evaluated.

To tackle this policy issue, in this study, tendencies and structures of personal criteria for the forms of research outputs are analyzed. Main findings are as follows;

- 1)Comparing between disciplines, on one hand, the forms of research outputs and the evaluation standards for them are far different. In addition, between positions the evaluation standards are found to be far different. These result suggests the difficulties of uniform quantitative evaluation over all attributions.
- 2)On the other hand, some common structure are found over disciplines. The features of the structure of the evaluation in each discipline are also described.
- 3)The evaluation of the forms of research outputs may affects the choice of the output forms he/she use.
- 4)In the field of engineering, domestic japanese articles may written not from academic point of view but from other views university includes.