

## 厳格な成績評価？

教養部解体・GP分布・公正

藤 村 正 司



# 厳格な成績評価？

教養部解体・GP分布・公正

藤村正司\*

「成績分布はかなり恣意的なものだから、みんなの成績も「C」は実質的には「B」と、また、「B」は「A」とそう変わらないのだ、などと決して言うてはならない。」(J.ローマン『大学のティーチング』316頁)

## 1. はじめに - 問題の設定

教養部の解体後、教養教育において何を教え、誰が担当するかが、10年を経た現在でも大きな問題でありつづけている。否、10年を経たからこそ、当初懸念された教養科目担当教官の補充の問題や新たな教養分野の模索が現実化しはじめたと言ってよい(寺崎：1999；有本編：2003)。

こうしたなか、平成10年に大学審議会は、教養部改組に伴う弊害を是正するために、「教養教育の重視」、「課題探求能力の育成」や「厳格な成績評価」といった授業内容・方法に踏み込んだ答申を打ち出した(大学審議会：1998)。言うまでもなく、授業内容や成績評価は、本来、科目担当教員の責任と信頼においてなされるべきものである。しかし、大学審が個々の教員レベルで実践すべき事柄にあえて自覚を促すのは、教員の授業運営や成績評価に対する不信があるからである。それは、「学部段階の教育については、……学期末の試験のみで成績評価が行われている、成績評価が甘く安易な進級・卒業認定が行われている、教養教育が軽視されているのではないかという危惧がある。」(大学審答申：1998, 11頁)と記されている通りである。

だが、このような大学審答申の、単位認定に関わる危惧と「厳格な成績評価」のススメは、どれだけ授業の運営や成績評価の実態を把握したものであろうか。また、そうした実践的な是正の方向性は個々の大学でどの程度浸透しているのか、また政策と実際の教員の意識や行為の間にどのようなズレが生まれるのかについては、実証的な検討を踏まえる必要がある。

そうした改革の影響を明確にする一つの指標として、本稿では成績評価に注目してみたい。ところで、これまでわが国では、初等中等教育と比べて、大学での成績評価について公的に議論されることは少なかった。その背景として入試選抜による学力保証説、大学の知識＝無効説、世間体説(授業料を支払った学生は落とせない)などがある。その一方で、入試科目数のスリム化と高校までの学習指導要領改訂による学力危機から、教育の質的保証が求められようになっている(民主教育協会：1999；2002)。だが、そもそもモデルとされるアメリカの大学と比べて、短大から大学への学生のモビリティの少ないわが国の高等教育においては、成績の標準化は重視されてこなかったし、何よりも成績は評価される側には決定的な影響を及ぼすが、評価する側には何の影響も及ぼさ

\* 新潟大学教育人間科学部教授

ない(Cahn : 1978, p.218)。いずれにせよ、成績評価にかかわる議論は、学力問題と裏腹の関係にありながら、教員の裁量と責任において自覚を促すという「信頼の論理」が支配的であった。

このような状況から、個々の大学において成績評価が科目別にどのようなになっているのか、それがマクロ的には教養部の解体やカリキュラムの再編でどう変化し、ミクロ的には単位認定の方法、担当者の所属、クラスサイズ、そして学生による授業評価の実践とどのように関係しているのか、具体的には「秀」と「優」、「可」と「不可」の境界をどう設定するかという悩ましい検討は回避され、教員の直感的判断に委ねられたままであった。

本稿では、新潟大学教養教育科目を事例にその成績分布を示し、成績分布に影響を与える諸要因を探り当てることを目的とする。成績評価は、外部者が異議をさしはさむことのできない不可侵な行為であり、また学生と教官を分かちもつとも非対称な領域である。あえて成績評価に注目するのは、評価の公正さを教員に求めることもあるが、それ以上に成績評価が学生による授業評価を構造化する視点を提供するとともに、何よりも教養部改組の効果、すなわち大学組織が弛緩するときどのような事態が生まれるのかを検証する時期になっていると考えるからである。

## 2. 予備的考察：日本とアメリカ

成績分布の分析に先立って、教員は学部学生の能力をどのように認識しているのか、また単位認定に関わる大学審答申の批判的な現状認識はどの程度実態を反映したものであるのか、授業方法におけるコミュニケーションのあり方はどのようなものか、これらの点から明らかにしておこう。ここでは、カーネギー大学教授職国際調査(1992年)から日本とアメリカを比較検討してみよう。

### 2-1. 教員からみた学部学生の能力

表1は、学部学生の能力について、教員はどのようにみているのか示したものである。比較のために日米ともに勤務大学を設置者と大学の水準によって4類型(国立・公立研究大学、国立・公立一般大学、私立研究大学、私立一般大学)に区分した<sup>1)</sup>。質問項目は【学生は十分な筆記及び口頭によるコミュニケーション能力を持っている】、【学生は十分な数学的・論理的能力を持っている】、【学生は何とかして単位を取れるだけの勉強しかしない】、【学生はよい成績を得るためにカンニングしようとする】、【現在の学生は5年前の学生よりもよく勉強する】である。表1から、日本とアメリカとでは、明らかに学部学生の質について教員の意見は異なっていることがわかる。アメリカのデータについては、ボイヤーが1980年代に実施されたカーネギー調査から、アメリカの学生が十分な学力をもっていないことを指摘しているが(ボイヤー：1988)、90年代初頭に実施された同調査についても同様である。そのことは、日本の教員の意見との比較によって鮮明になる。

例えば、【学生は十分な筆記及び口頭によるコミュニケーション能力を持っている】について、肯定的に回答する日本の教員が全体で30%に達するにときに、アメリカでは14.3%にすぎない。【学生は十分な数学的・論理的能力を持っている】とする教員は、日本が22.4%、アメリカは10.6%である。大学類型でみると、とくに日本の方が研究大学と一般大学の間に大きな開きがある。しか

し、日本の場合、問題なのは、学生は十分な能力をもっていると認識しつつも、【学生は何とかして単位を取れるだけの勉強しかしない】という意見を肯定する教員が全体で60%に及んでいることである(アメリカは34.4%)。

こうした学部学生の能力に対する彼我の認識の違いが、一般に、日本の大学では大学教育への関心が低く、アメリカで一般に高いとみなされる大きな理由の一つになっている。ここには治療教育の導入、教授法の開発、シラバスのあり方、必修カリキュラムの必要性の程度も含めて、苅谷がいうように日本とアメリカの入学選抜のしくみ、大学の格差構造、高校修了までに学生が身につける学力レベル、そして進路など成績の現実的な意味づけの違いに求められるのであろう(苅谷：1992)。それでは、このような学生の能力に対する教員の認識は、どのような授業実践と結びついているであろうか。以下では、単位認定の要件についてみてみよう。

表1 教員からみた学部学生の能力評価

	日本					有意水準	アメリカ					有意水準
	国立研究	国立一般	私立研究	私立一般	全体		公立研究	公立一般	私立研究	私立一般	全体	
a) 学生は十分な筆記・口頭によるコミュニケーション能力をもっている	47.9 〔284〕	27.6 〔519〕	60.9 〔69〕	23.2 〔838〕	30.1 〔1710〕	***	11.1 〔948〕	7.5 〔854〕	23.5 〔553〕	22.5 〔462〕	14.3 〔2817〕	***
b) 学生は十分な数学的・論理的能力をもっている	51.3 〔279〕	19.6 〔511〕	57.4 〔68〕	11.1 〔799〕	22.4 〔1657〕	***	10.3 〔825〕	5.8 〔738〕	17.8 〔449〕	12.1 〔380〕	10.6 〔2392〕	***
c) 学生はなんとか単位をとれるだけの勉強しかしない	51.8 〔280〕	65.6 〔520〕	46.4 〔69〕	61.7 〔846〕	60.6 〔1715〕	***	31.3 〔950〕	46.1 〔851〕	23.8 〔554〕	32.3 〔465〕	34.4 〔2820〕	***
d) 学生はよい成績を得るためにカンニングしようとする	15.9 〔264〕	21.2 〔505〕	17.4 〔69〕	16.2 〔780〕	17.7 〔1618〕	n.s.	12.6 〔926〕	30.1 〔845〕	11.9 〔544〕	13.3 〔435〕	18 〔2750〕	
e) 現在の学生は5年前の学生よりもよく勉強する	10.4 〔279〕	4 〔505〕	23.9 〔67〕	18.6 〔823〕	13 〔1674〕	***	9.9 〔881〕	15.2 〔802〕	12.8 〔514〕	11.6 〔398〕	12.4 〔2595〕	*

注：カーネギー大学教授職国際調査(1992)，有意水準 \* $p<5\%$ ，\*\*\* $p<0.1\%$  大学類型は本文参照。

## 2-2. 単位認定の要件

表2は、学部教養課程の授業を1科目以上担当する講師以上の常勤教員について、課程の単位認定にあたり、どのような要件を学生に課しているのかを示したものである。単位認定の要件は、【授業への出席状況】、【何回かの小論文の提出】、【主論文の提出】、【口頭で発表すること】、【クラス討議への積極的参加】、【試験を1回受けること】、【試験を2回以上受けること】、【特別な要件はない】である。表2より、すべての項目について日本よりもアメリカの教員の方が、能力の低いと評価される学生に対して単位認定に多くの要件を課す教員の比率が高いことがわかる。そのなかで日本と比較して、要件の違いがもっとも際だっているのは、【クラス討議への積極的参加】と【試験を2回以上受けること】である。【クラス討議への積極的参加】を課す教員は、全体で日本が14%にすぎないときに、アメリカの教員は60.1%にも及ぶ。また、【試験を2回以上受けること】を課す教員は、日本が42.8%、アメリカでは82.7%と2倍近くになる。【試験を1回受けること】に該当するアメリカの教員は、全体の8%にすぎない。

表2 学部教養課程の単位認定の要件

	日本						アメリカ					
	国立研究	国立一般	私立研究	私立一般	全体	有意水準	公立研究	公立一般	私立研究	私立一般	全体	有意水準
a) 授業への出席状況	51.6	60.5	59.5	73.9	66.4	***	65.7	76.1	57.6	84.2	75.3	***
b) 何回かの小論文の提出	28	26.5	35.1	35.5	31.9	n.s.	39.6	39.9	37.9	52.8	43.7	*
c) 主論文の提出	4.3	7.5	10.5	5.3	6.1	n.s.	14.7	19.9	197	20.1	19.2	n.s.
d) 口頭での発表	12.9	11	10.8	19.2	15.6	*	17.6	19.9	18.2	14.4	17.7	n.s.
e) クラス討議への積極的参加	10.8	13	13.5	15.5	14	n.s.	55.4	52	54	75.7	60.1	***
f) 試験を1回受けること	45.2	43	40.5	26.7	34.5	***	9.8	9.1	7.6	7	8.4	n.s.
g) 試験を2回以上受けること	26.9	29.5	37.8	54.4	42.8	***	83.2	84	80.3	81.4	82.7	n.s.
h) 特別な要件はない	6.5	2.5	5.4	3.2	3.5	n.s.	0	0	0	1.4	0.3	n.s.
N	93	200	37	375	705		102	306	66	215	689	

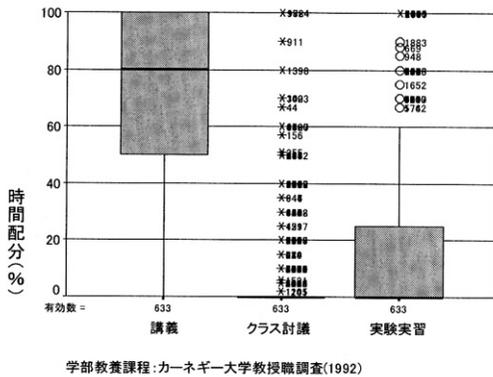
注：カーネギー大学教授職国際調査(1992)

だが、大学類型でみればバリエーションがある。日本では「私立一般大学」がアメリカ型に近いが、アメリカの「私立一般大学」(=教養カレッジ)ではさらに【何回かの小論文の提出】を課す教員が多い。他方、日本の国立大学は出席と1回の試験で単位認定の評価がなされる。また、学生の能力評価との関係で言えば、学生の能力評価の高い日本の研究大学ほど、単位認定要件は低くなっている。それだけ、研究大学では時間とエネルギーを教育に費やさなくてもすむのである。

要するに、90年代初頭の日本の教養教育における単位認定の要件とは、【授業への出席状況】と【1, 2回の試験】であるのに、アメリカでは【授業への出席状況】に加えて、【クラス討議への積極的参加】や【試験を2回以上受けること】など進級判定に手間暇をかけている。むろん、92年当時は Semester 制に移行する前であり、単純な比較はできないとしても、答申に記載される「成績評価基準は……学期末の試験のみでなく授業の出席状況、宿題への対応状況、レポート等の提出状況等、日常の学生の授業への取り組みと成果を考慮して多面的な設定をすることが望ましい」(大学審：49頁)という注文は、たしかに当たっているが、それは大学類型でいえば国立(研究)大学に向けて書かれた注文であると読める。もっとも、答申の「…学期末の試験のみでなく、授業の出席状況…」の箇所は、「……出席状況のみでなく、学期末の試験、……」が正しい。本稿の分析課題との関係でいえば、成績評価が改組前後でどう変化し、単位認定の要件によって成績評価のあり方がどのように変わるかである。

### 2-3. 授業方法

授業内容の改善についても触れておきたい。表2から授業方法にかかわる項目は、【クラス討議への積極的参加】のみである。答申ではとくに【クラス討議】についての記載はないが、日米の単位認定の要件を比較したとき、【クラス討議】のウエイトが大きく異なる点はみてきた通りである。それでは、教員は実際にどのくらいの時間を【クラス討議】に使っているのだろうか。図1と図2は、学部教養課程を担当する教員について、【講義】、【クラス討議】、【実験実習】、【その他】に費やす時間配分を100%としたときの、【講義】、【クラス討議】、【実験実習】の配分比率を描いたも



学部教養課程：カーネギー大学教授職調査(1992)

図1 授業方法の時間配分(%)：日本

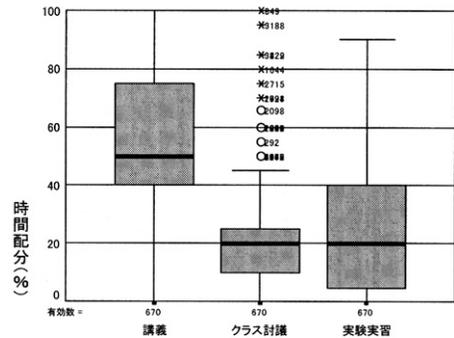


図2 授業方法の時間配分(%)：アメリカ

のである。これをみると、たしかに日本の大学の授業は【講義】中心で、【クラス討議】にはほとんど時間を割いていないことがわかる。対人関係が大きく介在しない講義中心であれば、たしかにTAの助けを必要としない。多大なエネルギーを要し、しかも効率の悪い【クラス討議】を行うアメリカの授業は、日本のモノローグ型・講壇型のコミュニケーション文化とアメリカのダイアローグ型・学習者支援型のスタイルの差異に求められるのかもしれない(荻谷：1992, 208頁)。

もっとも、アメリカの箱形図をみると、【クラス討議】の分布の中央値は20%程度である。日本のハズレ値だけの分布とは格段の差があるが、アメリカでも【講義】が授業の大半を占めていることにはかわりはないし、大学教師のすべてがそれほど【クラス討議】に熱心であるとはいえない。

しかし、荻谷も指摘するように、クラス討議に限らず、シラバスやTA、授業評価の導入が日本とアメリカの文化的差異を越えて、授業方法や成績評価のあり方を変える側面も軽視できない(荻谷：1992, 208頁)。以下、制度改革の創発的な側面について、成績分布を手がかりに探ってみよう。

### 3. 大綱化の衝撃

ここでは新潟大学の教養教育科目を事例に、教養部改組前後に教養科目担当教官にどのような変化が生じたのか検討し、次いで成績分布とその変動要因を探ってみよう。教養科目の成績評価を選んだのは、教養教育科目の管理化が全学の中でもっとも進んでいることによる。

#### 3-1. 教養科目担当教官の分散と集中

分析に先立って、表3によって新潟大学における教養部改組直前の教養科目とその担当者にどのような変化が生じたのか、また今後どのような事態が想定されるのか概観しておこう<sup>2)</sup>。教養部改組直前の1993年度現在、教養部教官数は98名。シラバスによると、この教養部教官が全科目の84%を担当していた。また、学部等教官のうち、教養教育を担当するものは172名(30%)であり、その半数は外国語科目を応援していたのである。

しかし、2002年度まで教養科目を担当する学部教官が全体の7割近くまで増加しているにもかか

表3 教養部改組前後の教養科目担当者数・担当時間数等の変遷

	教員現員数	教養科目担当者数	同担当時間数	一人当たり時間数
1993年度	学部等教官 574	172 ( 30% )	4,547	26.4
	教養部教官 98	98 ( 100% )	23,390	238.7
合 計	672	270 ( 40.2% )	27,937	103.5
1994年度	666	335 ( 50.3% )	22,897	68.4
2001年度	778	512 ( 65.8% )	21,066	41.1
2002年度	768	532 ( 69.3% )	20,873	39.2

注：助手は除く，時間数は外国人教師と非常勤講師は除く。 新潟大学学生部教務課管理係調べ

ならず，教養科目担当時間数は漸減している。教養教官担当者比率(括弧内)が増加したのは，全学出勤体制の浸透によるところが大きい，複数教官で1科目の授業を担当するようになったこともある。実際，1993年度のシラバスでは，非常勤も含めて一人で授業を担当する教養科目の割合が全体の92.7%であったのが，2001年度では80%まで減少している。こうした複数教官によるオムニバス形式が，授業のあり方や成績評価にどのような影響をもたらすかは実証の問題である。一方，担当時間数が減少しているのは，定員削減によって教養教育を担当していた教官の退職・異動後の未補充による。定員削減を被るのは教養教育担当の有無に関わらないが，英語と違って代替しにくい第二外国語科目では，定削による担当教官の未補充は，とくに地方大学では危機的な状況にあるといつてよい。

さて，2001年度のシラバスによれば，セメスター制の導入と全学出勤体制によって全教養科目数は改組前の776科目から1,095科目まで増加した。とくに著しく増加した科目群は，「総合科目群」と少人数制を導入した「外国語科目群」である。「総合科目群」は，1993年の31科目(全体の4%)から2001年には92科目(8.4%)，外国語科目群は277科目(35.7%)から543(50%)に増加した。果たして，元教養教官が担当する全科目の割合は，22.1%(242科目)まで半減したが，その242科目のうち，45.5%までが外国語科目群に集中しているのである(1993年度は26.4%)。もっとも，教養科目数は増加したが，卒業履修基準における教養科目単位の割合は，改組前の平成5年では最大42.4%，最小38.5%であったのが，2002年度には最大33.6%，最小24.6%までほぼ10%減少したのである。

こうして大綱化によって教養教育科目は増加し，全学に分散したように見えるが，教員一人当たりの時間数と教養科目の比重は軽くなった。しかし，そのなかで全学出勤が適わない外国語科目については元教養部教官にいつそう集中しているのである。今や教養科目の半数を占めるに至った，外国語科目担当者の確保が教養教育にかかわる緊要な課題になっているといつてよい。とくに，旧制高校時代に引き続き新制後50年の間，形骸化を指摘されつつも教養教育のシンボルとされた第二外国語科目が，担当者の未補充によってダメージをもっとも受けようとしているのである。それでは，成績分布が教養部改組前後でどのように変化したのか探ってみよう。

### 3-2. 甘くなった成績評価

図3と図4は、1993年度(平成5年)と2001年度(平成13年)に開講された全教養科目について成績(素点)を、5段階=「秀」(90点以上),「優」(80-89点),「良」(70-79点),「可」(60-69点),「不可(59点以下)・途中放棄」に換算し、受講生全体に占める5段階評価(比率)の分布を箱形図で描いたものである<sup>3)</sup>。二つの図を比べることで、教養部改組前と後では、成績分布に変化が生じていることがわかる。1993年の成績評価は正規分布であったのが、2001年度では相対的に「秀」と「優」の割合が増え、「可」と「不可・途中放棄」の比率が減少していることは明らかである。

「秀」と「優」の比率が増え、したがって「可」と「不可・放棄」が減少する理由は、四つ考えられる。第1は、優秀な学生が入学したこと。第2は、単位の意味が学生に周知されて、授業外の勉強時間が増えたこと。第3は、全学的なFDによって授業の改善効果が現れたこと。そして第4は、成績評価の基準が下がったことである。しかし、入試科目数の削減と学力低下が指摘される中、この8年間に資質の高い学生が新潟大学に入学したとは考えにくいし、第2の学生の勉強時間が増えたということも疑わしい。したがって、指導法の改善により教育効果が向上したか、評価する側が甘くなったかのいずれかと考えるのが妥当である。

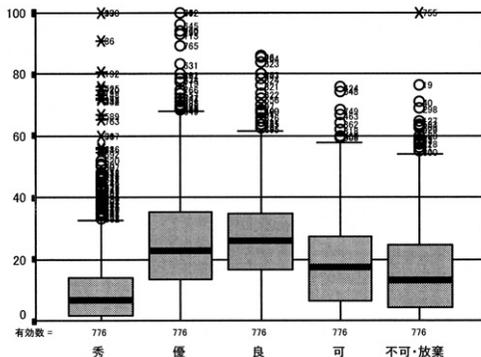


図3 教養科目GP分布：1993年度

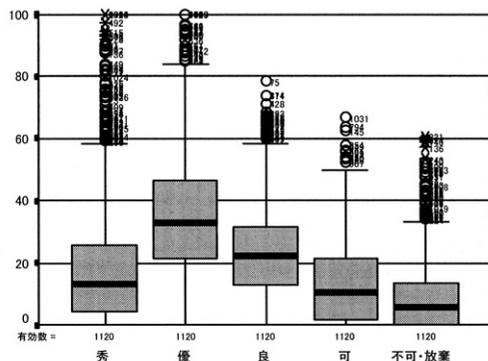


図4 教養科目GP分布：2001年度

もっとも、高い評価を与えても誰も傷つくわけではないから問題がないのかもしれないが、大学の成績評価に対して学生が不信感をもつのは、評価が不当に低い場合だけではない。評価が高すぎても不当だと感じる。それが同じ分野の科目間でばらつきが大きければ、公正の点から問題がないとはいえない。しかし、公正の観点から問題だといっても、果たしてどの程度のばらつきであれば不当であるといえるのか、その基準の設定は曖昧である。大学での成績評価は到達度評価のはずだが、「不可」を出さないよう配慮しているのが実際であろう。すなわち、平均に近づけるように評価基準を底上げ、もしくは底下げしているとすれば、自分では気づかないで相対評価をしていることになる。大学の成績評価が恣意的というのは、こういうことであろう。

それでは、もし教育効果が上がったのだとすれば、どのような教授法が功を奏したのか。あるいは評価基準が下がったのだとすれば、それは教養教育の再編とどう関わるのか。これについては、カリキュラム改革の対応、授業方法の改善状況と成績評価の関係から、以下のような仮説を設定し

てみたい。

- (1) マスプロ解消と少人数クラスの導入により教育効果があがった。
- (2) スタディスキルズ(学習法)など自学部生を指導する科目が共通科目に位置づけられたために、学生と教官の間に人間関係 = 「知っている学生」ができ、厳しい評価がつけにくくなった。改組前であれば「他人の学生」に対して厳しい評価をつけることもできたが、改組後は学部に進級してくる学生に対して厳しい評価をつけにくくなった。つまり、一年次生を学部学生とみなすようになったため、庇護移動が作動するようになった。
- (3) 全学出勤体制により、初めて教養教育を担当する教員が増加したが、負担軽減のために複数教官が担当するオムニバス形式が増えるとともに、評価基準が曖昧になり、総じて評価が甘くなった。
- (4) 教養科目を専門科目への「導入」と位置づけたため、総じて評価基準が下がった。
- (5) カリキュラムの区分が廃止されたことにより、環境、生命、共生、地域学など新たな教養分野の誕生とともに、個性化科目、コア科目、総合科目が多く立ち上がり、評価基準が曖昧になった。
- (6) 答申にしたがって、成績評価を厳格にするため、成績優秀者に対して新たに「秀」を設けたが、「秀」と「優」の区分について教員の間で理解が得られていないことが考えられる。そうだとすれば、厳格な成績評価の「意図せざる結果」だといえる。実際、1993年度では受講生の6割以上に「秀」( = 科目の目標を超えている ) をつけた気前のよい科目は全体の1.8%にすぎなかったのが、2001年度には5.6%と増加している。あるいは、履修単位の上限を設定したため、学生にとってリスク回避のために過剰な「保険単位」をとりにくくなったこと、また教員も厳しい評価をつけずらくなったことがある。

おそらく制度改革に伴い、教養教育をめぐって以上のような関係が複合して相対的に高い成績評価もたらしたのであろう。いずれにせよ、甘い評価によって学生の能力の実態が見えにくくなっているのである。しかし、果たして教官の所属、複数担当、学生との人間関係、クラスサイズなどによって、成績評価に偏りが現れるものであろうか。

### 3-3. 教養部改組と成績評価

表4は、教養改組前の1993年度と2001年度の教養教育全科目を合併したデータを用いて、2001年度交互採用ダミーを含む回帰分析の結果を示したものである。説明変数として、2001年度ダミーの他に、上記の仮説を考慮しつつ両年度を通じて利用可能な変数を投入した。すなわち、科目担当者の人数(複数担当 = 1, 一人担当 = 0), 教官の所属(元教養教官 = 1, 学部等教官 = 0), クラスサイズとして受講生数, そしてセメスター制導入前の1993年度については開講区分(通年 = 1, 半期 = 0)を投入している。なお、学部等教官には教養部経験者を含む。シラバスの記載内容から、2時点の比較ができるのはこれらの変数のみである。また、2001年度では「情報処理科目」が「総合科目」から独立したが、1993年度には「総合科目」に含まれていたため合併した。被説明変数は、煩雑さを避けるために図3と図4の秀と優、可と不可(途中放棄を含む)を加えた比率を用いる。

表4から、2001年度ダミーが「秀・優」評価についてプラス、「可・不可」ではマイナスの有意な効果をもつことがわかる。すなわち、他の変数を一定とすれば、「可・不可」の比率も同様に12%ほど低くなっている。クラスサイズを示す受講生数は、成績評価とマイナスの関係にあることがわかる。マスプロ授業ほど評価が低くなりやすく、少人数制になると評価が良くなる傾向がある。二つの理由が考えられる。一つは、受講数が多いほど、途中放棄する学生や「知らない学生」が多くなるから、厳しい評価をつけやすくなること、今一つは授業効果を媒介とするもので、少人数クラスほど学生の理解度が深まり、高い成績評価につながったとみることできる。次いで、教養教官ダミーは5%水準で有意ではないが、(元)教養教官ほど評価が厳しい(つまり、学部等教官はやさしい)傾向にある。そして、複数担当によるオムニバス形式は、一人で担当する場合よりも評価が甘くなっているのである。

このように学生の能力をどう評価するかは根元的問題に違いないが、教養部改組によって、さらに授業形式、教官の異動、クラスサイズ等によって能力評価は偏りをもってくるのである。ここで興味深いのは、科目群ダミーを投入すると、通年開講(4単位)の授業は符合条件が変わることである。つまり、科目群のバラツキを一定とすると、通年の授業では成績評価が高くなる。そうだとすれば、 Semester制の現在では、厳しい評価をつけても痛みは小さくなったといえるのかもしれない。通年4単位と半期2単位の重みの違いが成績評価に与えた効果は、大きかったのである。

表4 教養科目の成績評価に及ぼす影響(2001年度交互作用ダミーを含む)

	秀・優				可・不可			
	b		b		b		b	
受講者数	-.105d	-.265	-.098d	-.249	.074d	.226	.059d	.180
元教養教官	-.574	-.011	-1.280	-.024	-.041	-.001	-.041	-.001
教養教官*2001年度(d)	-.997	-.013	-3.586	-.048	2.019	.032	3.251a	.052
複数担当(d)	2.508	.035	3.377b	.047	-4.544d	-.076	-5.646d	-.094
2001年度(d)	12.199d	.235	9.163c	.177	-8.759d	-.202	-12.280d	-.284
通年開講(d)	-5.302c	-.083	9.405c	.147	6.452d	.121	-10.360d	-.194
総合科目			-9.020b	-.101			15.098d	.203
総合科目*2001年度(d)			7.153	.073			-11.109b	-.136
人文科学			-14.746d	-.169			13.101d	.180
人文科目*2001年度(d)			12.764c	.109			-10.562c	-.108
自然科学			-11.842d	-.171			7.567c	.131
自然科学2001年度(d)			5.166	.054			1.441	.018
外国語			-20.702d	-.403			17.365d	.405
外国語*2001年度(d)			19.835d	.352			-14.648d	-.312
保健体育			10.527c	.135			-20.479d	-.314
保健体育*2001年度(d)			11.591c	.103			5.749	.061
社会科学			(base)	(base)			(base)	(base)
決定係数	.166		.252		.138		.271	
N	1.889		1.889		1.889		1.889	

注：有意水準 a...p<10% , b...p<5% , c...p<1% , d...p<0.1%

なお、ここでは科目群間の成績評価のバラツキについては立ち入らないが、表4より科目群間で成績評価のバラツキも大きいことがわかる。公正の観点から言えば、科目群間の差異よりも科目群内では調整が必要であろうが、総じて責任の所在が曖昧になっているのである。

#### 3-4. 評価方法・授業評価・GP分布

成績評価については、近年のシラバスにおいては評価方法やその配分比が明記されるように改善されている。そこには、出席、試験、レポート、平常点、授業参加度などを総合して評価すると記載されている。シラバスのこうした契約的側面が、成績評価の「公正さ」を判断する根拠になるからである(苅谷：1992, 139頁)。また、近年、成績評価は授業の構成要素の一つとして捉えられている(吉田：2003)。例えば、学生による授業評価のなかに成績評価項目を取り入れ、学生の授業満足度と教員の成績評価の関連が検討されている(慶伊・飯島・木村：1994)。

ここでは、2001年度の成績データにシラバス(「教養科目講義概要」と科目ごとの授業評価を繋いだデータファイルを用いて、評価方法や授業評価が成績評価にどのような影響を与えているのか検討してみよう。シラバスに記載される成績判定基準とともに、授業評価が実際に成績評価に影響を及ぼすことが明らかになれば、授業の構成要素として一つとして成績評価を位置づけることの妥当性が得られるし、授業評価の利用法としても有益な情報を得ることができよう。

表5 評価方法と授業評価の追加(2001年度)

	外国語科目				総合・人文・社会・自然科学科目			
	秀・優		可・不可		秀・優		可・不可	
教員属性								
複数担当	6.088a	6.717a	- 4.829a	- 5.389b	- 3.719	- 6.199a	.556	1.863
元教養教官	- 6.018a	- 3.126	3.747	1.203	- 9.854	- 8.819b	1.782	4.035
同じ学部	- 2.364	- 4.020	.227	1.640	6.847b	7.578b	- 4.941b	- 3.667
評価方法								
試験	- 4.719	- 3.574	4.667	4.296	- 9.015b	- 3.755	15.317c	2.255
レポート	- 2.856	.233	2.462	- 1.035	4.571	3.950	1.646	- 5.211a
出席点	- .003	.117	- 3.074a	- 2.919a	1.048	.961	- .061	- 1.828
平常点	2.542	2.642	- 3.629a	- 3.673a	10.236c	- 1.042	- 9.326c	- 2.497
授業評価								
満足興味		4.479c		- 4.699a		4.428c		- 6.987c
予習復習		6.064c		- 2.961		4.039a		- .056
教室環境		1.023		- .820		1.550		- .540
受講生数	- .282c	- .122	.129b	.007	- .061c	- .018	.041c	.012
決定係数	.038	.068	.042	.083	.197	.146	.294	.196
N	542	530	542	530	444	353	444	353

注：有意水準 a... $p < 10\%$ , b... $p < 5\%$ , c... $p < 1\%$ , 教員属性と評価方法はダミー変数, 授業評価は因子得点(注4参加)

表5は、2001年度の成績評価を「秀・優」と「可・不可」の比率に区分したときの諸要因の効果をみたものである。分析にあたって、科目の半数を占める「外国語科目群」と選択科目を含む「総合科目・人文科学・社会科学・自然科学」（以下、一般科目）に区分した（実技系の保健体育科目群は除外）。説明変数は、表4で用いた変数に、【評価方法】（試験、レポート、出席点、平常点）の記載の有無、学生による【授業評価】の因子得点<sup>4</sup>、そして学生と教員の人間関係指標として【同じ学部】に所属するか否か加えた。所属する学部1年次学生を担当する教官は、そうでない場合と比べて評価にどのような影響を与えるのか検討するためである（仮説2）。

まず、決定係数の大きさから、「外国語科目群」の場合、一般科目群に比べてモデルの説明力は高くない。「外国語科目群」は必修科目であり、教授法も似ているからであろう。また、外国人非常勤講師が多いことからシラバスに記載される評価方法も厳格である。「外国語科目群」の記載率は、レポート（14.2%）と出席点（68.3%）を別にして、試験が95.4%、平常点が72.7%となっている（一般科目の記載率は、試験57.2%、レポート60.8%、出席点69.3%、平常点32.1%）。

教員の属性については、表4でみたように元教養教官は、学部教官よりも「秀・優」をつけることは少ないこと、つまり学部教官ほど評価が甘いことがわかる。しかも、学生と教官が同じ学部所属する科目の場合、そうした庇護的傾向が一般科目で顕著である。そうだとすれば、学部学生に対する甘い評価は、おそらく学部専門科目や大学院の授業ほど顕著であるといえるかもしれない。

評価方法についてはどうか。外国語科目群では、一般科目群と比べてどの評価方法も「秀・優」に結びつかない。むしろ、表6に見るように、外国語では「出席点」や「平常点」が「可・不可」に影響を及ぼしていることがわかる。一般科目群では、「秀・優」の評価は「試験」と「平常点」が効いている。たしかに、「試験」は学生の能力を点数化するから厳しい評価になりやすい。ここで興味深いのは「平常点」である。「平常点」がシラバスに記載される科目は、実際に成績評価の高低を決定する目安になっている。だが、レポートは、他の評価方法と比べて成績評価と必ずしも明確な関係がみられない。それは教室の内と外を繋ぐレポートは、評価方法のなかで恣意的・直感的であるからであろう。どのようにレポートを評価するかは検討を要しよう。

次いで、学生による授業評価から得られた3つの因子得点をモデルに投入すると、【授業満足因子】＝「授業内容は興味深い、総合的に満足」と【予習復習因子】＝「講義のために授業時間以外の学習（予習・復習）」をした科目や予備知識のある場合、成績評価とプラスの関係にあることがわかる。このことは、授業改善が成績評価にリンクするものであることを示すものである。しかし、3つの授業評価因子を投入すると、「秀・優」に及ぼす「評価方法」の有意な効果がみられなくなる。このことから、成績評価の方法が「秀・優」に及ぼす影響は直接的なものではなく、学生の授業に対する興味関心や授業外の勉強時間を媒介にしたものであることがわかる。

つまり、試験や平常点などの評価方法のあり方によって、授業への興味関心や授業外での学習の程度が異なり、それによってとくに予習復習を必要としない一般科目について成績評価に影響が現れたのである。さらに、授業評価を一定とすると、一般科目では教員の属性がすべて有意に転じる。講義内容の興味関心や授業外の学習時間が同じ程度であれば、仮説2で示したように、評価が高くなるのは、受講生と同じ学部所属する教員の一般科目であることがわかる。

ところが、「可・不可」については、授業評価を一定としても評価方法は有意である。とくに、出席点、レポート、平常点は、「可・不可」の救済装置として働いていることがわかる。最後に、クラスサイズを表す「受講生数」をみると表4でみたように、受講生が少ないほど評価が高くなる傾向にある。とくに、外国語科目については、一般科目に比べて3倍から4倍ほど「秀・優」の比率を高める効果がある。複数担当も含めて少人数制の導入の効果は、外国語修得の効率を高めていると言えるが、それだけコストを要するともいえるのである。

#### 4. おわりに

学力や能力という本質的に定性的なものを、ABCにランクづけするのは、多少ともジレンマを伴うものである。それを量的に評価するというのは、至難の業である。たしかに、「正しい答えとまちがった答えの数は計算できるが、その学生が「A」の才能をもっているのか、「B」の才能を持っているのか、あるいはまた、他にどんな種類のすぐれた才能を持っているかを決定するのは、教師の直感的判断だけである」(マサチューセッツ工科大学：1974, 35頁)。

しかし、「厳格な成績評価」という注文をつけた大学審の答申は、単位認定にかかわってルーズな学部教育への不信と改善を表明したものだが、それはアメリカとの比較でも裏付けることができた。さらに、我々の成績データは、教養部の改組後に成績評価が相対的に甘くなるという結果を示した。個性化科目、コア科目、四年一貫教育、共通教育科目の名のもとで、学生の学力低下問題とは裏腹に、学生に甘い評価を与えるようになっていたのである。これは一つには、答申にしたがって成績評価を厳格にするために、新たに「秀」(90点以上)を設けたが、担当教員の間で「秀」の成績評価に統一がとれていなかったことがある。これまでの「優」が「秀」に、「良」が「優」と変わらなくなった。学生に対して履修科目登録の上限を設定するのであれば、教員に対して成績評価の上限設定が指導されてしかるべきであろう。

実際、成績評価を底上げしたのは、教養部改組後、教養教育に駆り出された学部教員によるところが大きい。とは言っても、誰が教員に成績の付け方を指導するのであろうか。大綱化後の10年は、少なくない国立大学で大学院の重点化や教員養成系学部の改組が進んだ10年でもあった。教養部改組は大学院の重点化や教育学部の改組に人的資源を抛出したが、教養教育については全学出動体制の掛け声のもとに、現実的には全学無責任体制が進行した。新たに立ち上がった大学教育研究等センターと学部から選出された元教養部教官からなる調整機関では、責任の所在が明確になりにくいのである。

もっとも、教員がまったく教養教育に無責任であったのではない。それどころか、成績評価は、教員の所属のみならず、シラバスに記載される成績評価の方法や授業の満足さの影響も受けている。教養部という「大道具」の喪失は、新たな「小道具」によって、つまり教員の授業改善によって補われているのであり、成績評価は授業評価とともに授業を構成する要素として捉えることが実証された。しかし、そうした授業改善や評価方法とは無関係に、教養部改組が教養教育に及ぼした衝撃は大きかった。本稿では、成績評価の分布に着目したが、法人化後は、教員組織の系列化、カリキ

ユラムのスリム化，ないしはワークシェアリングによって学部間の敷居が低くなる。組織の凝集性が低下するときに，大学教育の質的保証がどのように保たれるのかは今後の検討課題である。

以上の分析結果は，一国立大学の事例である。むろん，成績データを正規分布に近づけることに夢中になる必要もその余裕もない。だが，GPAの運用にともない，成績の評価基準を明確にしておくことは，公正の観点からも重要である。各大学で個別に検証されるべき事柄であると考えられる。

付記：本稿は，平成14年度新潟大学教養教育実施委員会経費による成果の一部である。教養教育科目成績データのGP変換にあたっては，新潟大学学生部教務課管理係の協力を得た。記して謝意を表したい。

### 【注】

- 1) アメリカの大学類型の区分は，カーネギー分類の「研究大学」と「大学院大学」を【研究大学】，「総合大学」と「教養カレッジ大学」を【一般大学】とした（江原：2003，70-71頁）。
- 2) 教養部の構造特性と改組による教員の移動，授業担当数，カリキュラム改革の全国的な動向については，吉田(2002a；2002b)，山野井(2003)，冠野(2003)参照。
- 3) 科目群別の成績分布は，藤村(2002)参照。
- 4) 授業評価アンケートは，全科目について授業の最後に，したがって成績評価の前に実施されるものである（新潟大学大学教育開発研究センター『平成13年度 学生による授業アンケート調査結果データ』）。18項目からなる変数は，各質問項目の4段階加算平均値に基づいている。因子分析により3因子（累積寄与率70%）が抽出された。3因子に対応する質問項目と因子負荷量は，第1因子【授業の興味深さ】：「講義内容は興味深かった」(0.879)，「講義内容に総合的満足している」(0.854)，「講義は後輩・知人に勧めたい」(0.818)，「講義の内容・説明はわかりやすい」(0.814)，第2因子【学生の予習・復習】：「授業時間外に予習・復習をした」(0.789)，「受講前に講義内容の予備知識があった」(0.766)，講義は学生の反応を見ながら進行していた」(0.742)，第3因子【教室の環境】：「講義室の状態(広さ・明るさ)は適切だった」(0.825)，「受講生数は適切だった」(0.696)である。

### 【参考文献】

天野郁夫『大学に教育革命を』有信堂，1997。

有本章編『大学のカリキュラム改革』玉川大学出版部，2003。

江原武一「大学教員のみた日米の大学 90年代初頭」『京都大学大学院教育学研究科紀要』第49号，2003，69-91頁。

荻谷綱彦『アメリカの大学 ニッポンの大学』玉川大学出版部，1992。

冠野文「教員の意識とカリキュラム改革 教養部解体がもたらしたもの」有本編，前掲書，44-59頁。

- 慶伊富長・飯島泰蔵・木村正行「北陸先端科学技術大学の自己評価」『高等教育研究紀要』第14号，1994.
- 国立大学協会・教養教育に関する特別委員会『教養教育の改善に関する調査報告書』1995.
- 大学審議会『21世紀の大学像と今後の改革方策について』1998.
- 寺崎昌男『大学教育の創造』東信堂，1999.
- 藤村正司「新潟大学教養教育GPデータにみる成績評価の分布」『大学教育研究年報』（新潟大学大学教育開発研究センター）第8号，2003.
- ポイヤール，E.，喜多村和之・館昭・伊藤彰浩訳『アメリカの大学・カレッジ』リクルート出版，1988.
- マサチューセッツ工科大学『教師と学生』IDE教育資料，第44集，1974.
- 民主教育協会『厳格な成績評価』IDE，No.405，1999.
- 民主教育協会『成績評価と質の保証』IDE，No.449，2002.
- J.ローマン，阿部美哉監訳『大学のティーチング』玉川大学出版部，1987.
- 山野井敦徳「教養教育と実施組織」有本編，前掲書，105-119頁.
- 吉田文「教養部の形成と解体 教員の配属の視点から」『国立学校財務センター研究報告』第6号，2002a.
- 吉田文「わが国におけるFDをめぐる10年と今後の展開」三尾忠男・吉田文編『FDが大学教育を変える』文葉社，2003.
- Cahn,S. (ed), *Scholars Who Teach*, Nelson-Hall, 1978.
- Yoshida, A., 'The curriculum reforms of the 1990s: What has changed?' *Higher Education*, Vol.43, 2002b, pp.43-63.

## What Makes Marks Liberally? : Reorganization of the Faculty of Liberal Arts, Distribution of Grade Points, and Equity

Masashi FUJIMURA\*

The purpose of this paper is to analyze the effects of the reform of undergraduate education, which has been taking place following relaxation of the regulative standard for universities in 1991. Since then, many universities have started to reorganize their faculty of liberal arts, to create a new curriculum, to introduce teaching assessment by students, to publish syllabuses, and to make use of GPA. But, there has been no attempt to evaluate the relationship between macro level reform and such micro level efforts at each university.

In order to evaluate the effects of such a de-regulative policy, we use grade point data sets of Niigata University, which were collected to combine the ratios of grade points, A, B, C and D on each subject, evaluation methods by teachers and teaching assessment by students in two periods: one was 1993 and the other in 2001. In 1994, general education was reorganized at Niigata University. Using these data sets, evaluation of grade points by teachers has been compared between the two periods, and the influences of class size, faculty affiliation, and methods of evaluation (exam, homework/report, class participation, class attendance), with the interactive effect of a 2001-year-dummy on the ratio of GPA examined. Box-plots are used to compare the allocation of GPA's in each period, and regression analyses are conducted.

The results show that grade points have been marked liberally over 8 years, and that the distribution of grade points is explained by the 2001-year-dummy and interactive 2001-year-dummy, after controlling for other variables.

Based on these findings, we argue that generous marks arise from the reorganization of the department of liberal arts. Of course, innovations and efforts to improve the method of teaching also contribute to the good marks by students. But, under strong pressure to provide research-oriented reform, in parallel with the relaxation of the regulative standard, the realities of student ability were concealed by assuming a logic of confidence in the evaluation of grades given by teachers.

---

\* Professor, Faculty of Education and Human Sciences, Niigata University