

原 著

授業評価と成績に基づいた大学カリキュラムのマッピング(I)： 授業の諸特性による影響

木村堅一・佐久本功達

Appraising the Curriculum Through Two-dimensional Display of Class Ratings and Student Grades and the Effects of Certain Class Characteristics

Kenichi Kimura and Kohtatsu Skakumoto

要 約

学生による授業評価は、大学教育の質を保証するものとして急速に普及してきた。しかし、授業評価はその手段の1つでしかなく、重要な指標である成績（学習成果）が軽視されつつある。本研究では、ある大学で1年間に開講された全ての授業科目を、成績と授業評価の2次元上に布置することで、大学カリキュラムの実態を的確に理解し、両次元で高水準を達成している授業科目の諸特性を抽出しようと試みた。分析の結果、科目区分、必修・選択、クラスサイズ、曜日、時限、授業形態など、教員の個人的努力では変えられない授業固有の諸特性によって、成績および授業評価の値が変化していた。最後に、授業評価研究における本研究の有効性とその限界について議論した。

キーワード： 教育評価、授業評価、成績、授業の特性、大学

Abstract

Class ratings have rapidly become a popular method of ensuring the quality of university education. However, they are merely one way of achieving this goal, at the expense of placing less emphasis on the important indicator of student grades (academic achievement). In this study, student grades and class ratings are compared for all classes offered at a university over a period of one year. The results are used to gain an accurate picture of the university's curriculum, including revealing the various attributes of classes for which both ratings of teaching and grades were high. Analysis showed that both student grades and class ratings depended on specific characteristics of the classes unrelated to the efforts of teaching staff, such as the subject, whether it was compulsory or elective, class size, day of class, time slot and class format. We also discuss the effectiveness and limitations of this study with regards to students' ratings of classes.

Key Words; educational assessment, class rating, grade, class characteristic, university

問題と目的

近年、大学では教育の質を保証することが最優先課題となっている。その背景には、大学の大量化と多世代化、学問の高度化と学際化、学力低下などの外部環境の変化

があり、従来の教育システムでの対処が難しくなってきたことが挙げられる。また、そもそも大学に内在する教育軽視の価値観も背景にあると考えられる。大学に特化した教育評価の有効な手段がなかったため、教員は教育活動より研究活動を優先し、マスメディアも卒業時より

名桜大学国際学部経営情報学科

〒905-8585 名護市字為又1220-1 E-mail: k.kimura@mail.meio-u.ac.jp k.sakumoto@mail.meio-u.ac.jp

Department of Management and Information Sciences, Meio University, 1220-1 Bimata, Nago, Okinawa 905-8585, Japan

も入学時の学力を報道した。さらに、研究組織の維持を優先させたため、教育カリキュラム自体が硬直化・肥大化し、時代のニーズに合わせたカリキュラム編成を弾力的かつ柔軟に行うことができなくなり、大学の卒業生に対する社会的期待は著しく低下してきた。大学に冬の時代が到来する中、自らが教育活動を点検・評価し、効果的に教育の質を保証する組織・システム開発を行うことは、時代の流れとして当然の帰結なのである。

この大学教育の質を高める取り組みを、ファカルティ・ディベロップメント (Faculty Development, 以後FDと省略) と呼ぶ。このFDを推進する上で重要な活動の1つに授業評価 (class rating) がある。多くの場合、授業終了時に受講生によって行われる授業の総括評価を指している。当論文で特に断りがない限り、「授業評価」は「受講生による授業評価」を意味する。授業評価の主な目的は、教員単位での授業改善、組織単位での授業改善、教員の人事評価などが挙げられる。

この授業評価に関する研究は2つに大別できる。1つは、カリキュラム編成に授業評価を活用することを目的とした研究である。例えば、授業評価の効率的な導入、カリキュラム改革での授業評価と他のFD活動との連携などである。櫻井 (1998) は授業評価の単独実施では教育改善が期待できないことを指摘しており、小笠原ら (2002) も授業評価の大学教育における位置づけを明確にし、特にシラバスを向上した上で、授業評価を活用する必要性を強調している。村田 (2004) は、Web上でのシラバスと連携した授業評価システムの開発を提案している。渡辺 (2001) は、授業評価の意義と限界を取り上げ、授業評価が授業改善の手段ではなく、目的化する危険性に警告を鳴らした。

もう1つの授業評価研究の方向性は、授業評価の規定因を解明することを目的としたものである。八ッ橋 (2000) は、授業評価(授業満足度)を高めるためには「役立つ可能性の高い内容の授業を学生が理解できるように行う」(p.174) ことが重要であると指摘している。浦上・林・石田 (1999) は、学生の受講動機と授業評価の関連性を検討し、受講動機によって学生の授業に対する意識や態度、またそれらの関連性が異なることを明らかにし、興味・関心に応じた授業選択を可能にする重要性を主張している。

松田ら (1999) は、受講生から授業評価、自己評価、授業選択態度、成績のデータを収集し、受講生が女性で、授業に対する興味・関心が高いことが良い授業の特性と相互作用し、真面目で積極的な学習態度を形成、その結果、学習達成感の高い学生が良い成績を得られそうな場合に高い授業評価を行うと報告している。牧野 (2001c) は、授業の内容評価は授業満足度と成績を促進し、教員評価は授業満足度のみ促進することを明らかにしている。

牧野 (2002) は、成績と授業評価に強い関連性は認められないが、成績と授業満足度の間に正の関連性があることを報告している。牧野 (2005) は、出席率と授業評価が正の関連性をもつことを指摘している。南 (2004) は、受講生数が授業評価に及ぼす影響を292クラスで分析し、クラスサイズが大きくなるほど授業評価が低下する傾向を一貫して見出し、授業評価は外部要因によって補正されるべきだと主張している。

以上のように、授業評価研究には、単に授業評価を改善する側面を検討するだけでなく、授業評価が教員によって統制できる内部要因 (教員の情熱、対話、板書、教材の工夫等)、あるいは受講生の特性 (受講態度等) や授業の特性 (時間割、授業制度等) などの外部要因によって影響を受けることを実証する「授業評価の信憑性」を問題視するための研究が存在している。

ところで、授業評価以外の重要な授業改善効果の指標に「成績 (学習成果)」がある。これまで、授業評価と成績との関連性をみた研究は存在するが、一部の授業における学生の成績と授業評価との関連性を検討したものが多く。安岡ら (1989) は、よく勉強した学生は授業を高く評価するため、学生の成績と授業評価は正の関連性をもつと予測したが、支持されなかった。牧野 (2001a, 2001b) は、実際の成績ではなく、主観的な学習成果は授業評価と関連することを報告している。なお、授業評価は扱っていないが、平均成績が、科目区分や受講生数などの外部要因によって影響を受けることを西山 (2005) が報告している。

授業評価の向上は学生の視点から授業改善を促進する指標であるのに対し、成績の向上は教員の視点から授業改善を促進する指標と位置づけられる。もちろん、この2つの評価次元は独立していると考えられる。例えば、授業評価が低ければ、教員の態度、授業内容、授業方法、評価方法などに問題が生じている可能性がある。しかし、そのような授業でも学習成果が保証されている授業は存在するであろう。反対に授業評価が高くとも学習成果 (例、授業目標の達成度) が保証されていない授業も存在すると考えられる。学習成果が低ければ、授業形態や教授方法を改善したり、あるいは前提となる学習基準を明確にし、受講生を選抜したりするなどの処置が必要となってくるだろう。その場合、学習成果を無視し、学生による授業評価だけに依拠するのは問題となる場合もあるだろう。

そのため、成績 (学習成果) は本来、良い授業の重要な指標であるべきであろう。しかし、授業評価研究が必ずしも成績を同時に扱ってきたとはいえない。他方、授業評価と成績を同時に取り上げた研究の多くが、単一あるいは少数の授業科目を対象としていたため、カリキュラム全体における授業評価と成績の両次元から、個々の

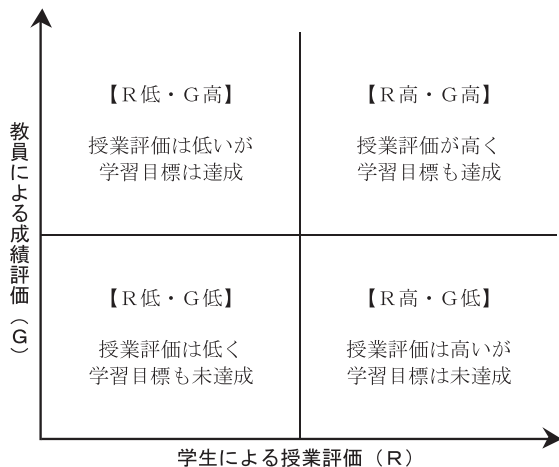


図1. 大学のカリキュラムにおける授業評価と成績に基づく2次元的理解。

授業を相対的に評価するといった視点が脱落してきたと言える。

本研究では、大学全体のカリキュラムを構成する科目を、それぞれの成績と授業評価に基づいた2次元上に布置することで、従来よりも利用しやすく、かつ適切な授業改善の理解を促進する図式を提案することを目的とする。同時に、授業固有の特性（外部要因）が成績と授業評価に及ぼす影響度を検討することで、授業評価と成績に基づき、教員個人の教育活動を評価する際の問題点についても明らかにする。

なお、ここで授業評価と成績の2次元によりカリキュラム（あるいは学生個人）を評価する意味について触れておく。授業評価(R)と成績(G)に基づく2次元のモデルを図1に示した。それぞれの次元を2群で折半し、4つの位相から大学カリキュラム（あるいは受講生）を捉えることが可能となる。なお、大学カリキュラムを評価する場合は、授業ごとに算出された平均成績と授業評価平均を用いる。

R高・G高 第1象限に布置されたデータは、授業評価も高く、教員が要求した学習目標も達成している科目、あるいは学生である。この象限に入る科目と学生に対して、教員は現状の良好な状態を維持することが将来の重要な目標となる。授業改善の緊急性は低い。

R低・G高 第2象限に布置されたデータは、授業評価は低いですが、教員が要求した学習目標を達成している科目、あるいは学生である。この象限に属する学生や科目に対しては、授業評価を低下させている原因（例、受講生が教え方、あるいは授業内容に共感できない等）を発見・解決すると同時に、教員が設定した学習目標の難易度が適切であるかを判断することが必要である。学習目標が低すぎた場合、授業評価は低くなるにもかかわらず、

学生は学習目標を達成しやすいと考えられる。

R高・G低 第4象限に布置されたデータは、授業評価は高いが、教員が要求する学習目標が未達成の科目、あるいは学生である。学生は授業内容や教授法に対して満足しているが、実際に成績評価を受ける段階で、教員が要求した学習基準をクリアできていない状況が伺える。授業内容と成績基準の不一致が考えられ、学生の学習意欲を維持・増進させ、学生の時間外学習活動を支援すると同時に、教員側も現在の授業評価を維持しつつ、学習目標達成を第一に考えた授業改善に取り組む必要がある。

R低・G低 第3象限に布置されたデータは、授業評価が低く、教員が要求する学習目標も未達成の科目、あるいは学生である。受講生にとって学習目標が高く設定されており、授業内容や授業方法がその高い学習目標を達成するために機能しておらず、教員と受講生が相互に低い評価を行う悪循環が生じている。授業改善の緊急性が最も高い領域である。

方法

対象となった授業科目 名桜大学で平成15年度前学期および後学期に開講された611の授業科目を研究対象とした。

データ利用の許可 受講生による授業評価および教員による成績評価に関するデータは、名桜大学自己点検・評価委員会が実施した「授業評価に関する意識調査(平成15年度・前学期/後学期)」の度数分布データ、ならびに平成15年度成績の度数分布データを授業名でマッチングできる形で入手した。これらのデータは、名桜大学の教育活動を評価・改善することを目的とし、教員および学生の個人情報特定できない統計処理を施すことを条件として、名桜大学自己点検・評価委員会から利用許可を得たものであった。

授業評価 授業評価21項目のうち、今回の分析では「総合的な授業評価」と「学生による成績予測」の2項目を利用した。具体的には、総合的な授業評価は「この授業の総合評価をして下さい」との質問に対して「5. 非常によい」「4. わりとよい」「3. どちらともいえない」「2. やや悪い」「1. 非常に悪い」で回答する5段階評価であった。成績予測は「この授業科目ではどの評価がとれると思いますか」という質問に「5. 秀」「4. 優」「3. 良」「2. 可」「1. 不可」の5段階で評価するものであった。授業科目ごとに授業評価平均 ($M=4.08, SD=.50$) と成績予測平均 ($M=3.10, SD=.48$) を算出した。なお、通年科目の授業評価データでは前学期のものは破棄し、後学期のものを採用した。以上の処理の結果、575科目の授業評価データが得られた。

成績 成績は「秀」を5、「優」を4、「良」を3、「可」

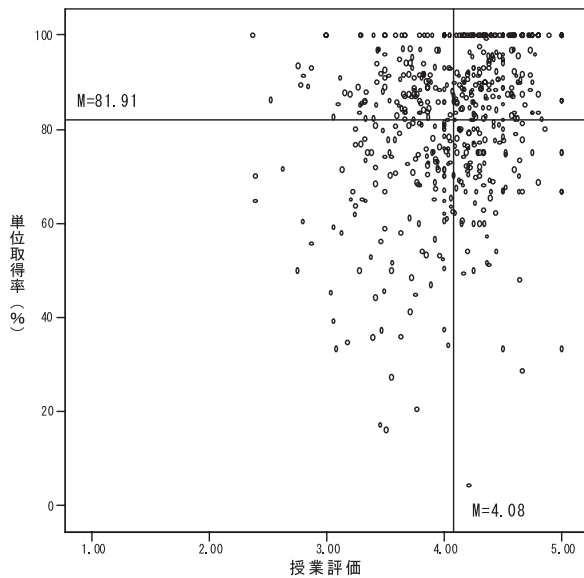


図2-1. 授業評価と単位取得率

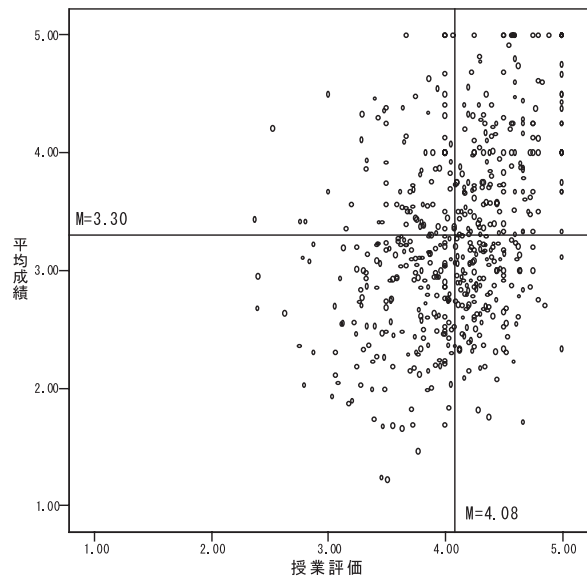


図2-2. 授業評価と平均成績

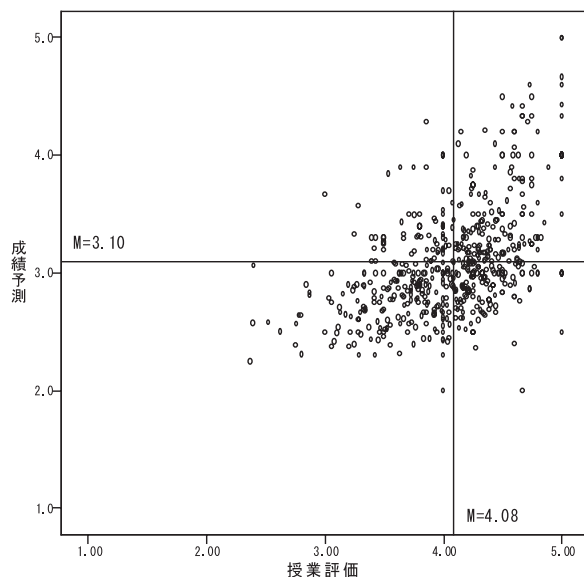


図2-3. 授業評価と成績予測

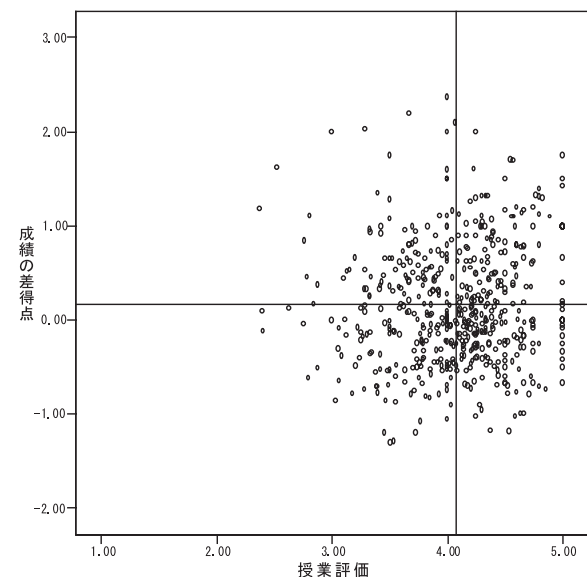


図2-4. 授業評価と成績差得点

を2、「不可」を1と数値化し、授業ごとに平均成績を算出した ($M=3.30$, $SD=.81$)。また、単位取得者数(「可」以上の成績取得者数)を受講生数で割り、単位取得率を算出した ($M=81.91\%$, $SD=16.69$)。なお、成績が「保留」であった場合、平均成績の算出ではそのサンプルを排除し、単位取得率の算出では単位未取得者と同等に扱った。最終的に、584科目の成績データが得られた。

授業の諸特性 教員個人では統制できない授業固有の特性と考えられる情報を、名桜大学の教務資料をもとに符号化した。611の授業科目の内訳は次の通りであった。

1. 授業形態 (講義482, 講読19, 演習84, 実習26)
2. 科目区分 (英語71, 第二外国語52, 教養51, 情報25, 保健体育11, 総合27, 国際文化113, 経営情報126,

観光産業99, 教職36)

3. 必修・選択指定 (必修200, 選択411)
4. 受講生数 (最少1, 最大204, 中央値29) なお, 30人未満を少人数クラス(287科目)に, 30人以上を大人数クラス(286科目)に分類した。
5. 曜日(月130, 火166, 水83, 木123, 金71, 週2回38) なお, 週2回ある授業科目については曜日を扱う分析では除外した。
6. 時限 (1時限106, 2時限125, 3時限123, 4時限121, 5時限88, 6時限10, 週2回38) なお, 週2回ある授業科目については時限を扱う分析では除外した。
7. 対象年次 (1年次246, 2年次159, 3年次159, 4年次47)

表 1. 授業形態による授業評価と平均成績の平均値と標準偏差

		講義	講読	演習	実習	F値
授業評価	M	4.05 ^a	4.14 ^a	4.27 ^a	4.26 ^a	$F(3, 545)=3.38$
	SD	0.5	0.52	0.5	0.32	$p<.05$
	N	467	19	77	12	
平均成績	M	3.1 ^a	3.34 ^a	4.27 ^b	4.11 ^b	$F(3, 545)=65.14$
	SD	0.7	0.52	0.66	0.64	$p<.001$
	N	466	19	74	25	

注) アルファベットが異なる平均値の間には $P<.05$ で有意差がある。

結 果

受講生による授業評価と成績評価の独立性を明らかにするために、横軸に授業評価の平均、縦軸に成績指標(単位取得率、平均成績、成績予測)をとった散布図を作成した(図2-1~図2-3)。なお、補足分析として平均成績から成績予測の差得点を算出し($M=.16$, $SD=.64$)、授業評価との関連性を検討した(図2-4)。

全体的傾向 単位取得率が高い授業ほど、授業評価は高いのであろうか。図2-1の散布図において、授業評価と単位取得率との間に弱い正の相関が認められた($r=.254$, $p<.001$)。つまり、授業評価が高い授業ほど、単位取得率は高く、逆に単位取得率が高い授業ほど、授業評価が高いという傾向が認められた。

良い成績を与える授業ほど、授業評価は高いのであろうか。図2-2の散布図の授業評価と平均成績との間にも弱い正の相関が認められた($r=.366$, $p<.001$)。つまり、授業評価が高い授業ほど、良い成績が与えられており、逆に良い成績が与えられる授業ほど、授業評価が高いという傾向が認められた。

良い成績が見込める授業ほど、授業評価は高いのであろうか。図2-3の散布図において、授業評価と成績予測の間には中程度の正の相関が認められた($r=.520$, $p<.001$)。つまり、授業評価が高い授業ほど、良い成績を受講生が見込んでおり、逆に受講生が良い成績を期待している授業ほど、授業評価が高いという傾向が認められた。

受講生が予想する成績よりも実際の成績が良い授業ほど、授業評価は高いのであろうか。実際の成績と成績予測とのギャップを算出し、この差得点で正の得点は「予想よりも良い成績がとれた」ことを、負の得点は「予想よりも悪い成績であった」ことを意味する。成績の差得点と授業評価の散布図が図2-4である。相関分析の結果、授業評価との間の相関は有意ではなかった($r=-.057$, $n.s.$)。なお、成績の差得点は $M=.16$, $SD=.64$ であり、平均成績と成績予測との間に正の有意な相関($r=.570$, $p<.001$) が認められたため、受講生は実際の成績をある程度正確に予想していたことが分かる。

もし、授業評価が成績(学習成果)と強い関連性をもつ

ていれば、授業評価を高める工夫をすることによって、成績にも良い影響が出ると予想される。一方で、それほど強い関連性が認められない場合、授業評価と成績はあくまで独立した指標であると考え、カリキュラム全体を授業評価だけでなく、成績指標を考慮した2次元上で表現する必要があるであろう。分析の結果、授業評価と成績は、全ての指標で正の相関関係が認められた。しかし、それは強い相関ではなく、学生の成績予測を用いた場合のみ中程度の相関が認められた。このため、授業評価だけに基づき、その授業の良否を判断することは大きな問題があることが示唆された。以後、成績指標の中で最も学習成果を敏感に反映していると考えられる「平均成績」を用いて分析を進める。

授業特性の影響 教員によっては統制が困難である授業固有の特性(科目区分、授業形態、受講生数など)が、授業評価と平均成績にどのような影響を及ぼすかを検討するため、次の3視点から分析を行った。第1に、授業の特性のグループごとに授業評価(あるいは平均成績)の平均値を算出し、その平均値に差が存在するのかを確認する。第2に、授業評価と平均成績の2次元上の座標上に全ての授業特性を布置し、大学全体のカリキュラムに対する授業固有の影響度を評価する。最後に、授業評価(あるいは平均成績)に対する表面的な影響度ではなく、各授業特性が独立して持つ影響度の強さを、カテゴリカル回帰分析を用いて評価を行う。

授業形態によって授業評価、あるいは平均成績の得点が異なるかを、一元配置分散分析を用いて検討した。表1の通り、授業評価に対して有意な主効果が認められたが($F(3, 545)=3.38$, $p<.05$)、授業形態4つの平均値を多重比較したところ、有意な差は認められなかった。平均成績に対して有意な主効果が認められ($F(3, 545)=65.14$, $p<.001$)、多重比較の結果、講義($M=3.10$)や講読(3.34)よりも実習(4.11)と演習(4.27)で平均成績の値が高かった。

科目区分の分散分析結果は表2に掲載した。授業評価($F(9, 539)=9.66$, $p<.001$)と平均成績($F(9, 539)=9.55$, $p<.001$)の両方に有意な主効果が認められ、授業評価では、第1に第二外国語($M=4.48$)の値が最も高く、第2

表 2 . 科目区分による授業評価と平均成績の平均値と標準偏差

	英語	第二外国語	教養基礎	情報	保健体育	総合	国際文化	経営情報	観光産業	教職	F値	
授業 評価	M	4.21 ^{bc}	4.48 ^c	3.82 ^a	3.85 ^a	4.22 ^{bc}	3.8 ^a	4.12 ^{ab}	4 ^{ab}	4.03 ^{ab}	4.3 ^{bc}	$F(9, 539)=9.66$
	SD	0.35	0.43	0.35	0.58	0.33	0.29	0.55	0.55	0.4	0.55	$p<.001$
	N	69	51	50	24	11	26	105	123	83	33	
平均 成績	M	3.06 ^{bc}	3.51 ^{cd}	2.81 ^{ab}	3.29 ^{bcd}	3.89 ^e	2.52 ^a	3.4 ^{cde}	3.36 ^{cde}	3.44 ^{cd}	3.69 ^{de}	$F(9, 539)=9.55$
	SD	0.71	0.88	0.73	0.54	0.69	0.55	0.85	0.76	0.78	0.67	$p<.001$
	N	65	49	50	25	11	27	103	123	95	36	

注) アルファベットが異なる平均値の間には $p<.05$ で有意差がある。

表 3 . 必修・選択による授業評価と平均成績の平均値と標準偏差

	必修	選択	F値	
授業 評価	M	4.08	4.09	$F(1, 547)=.44$
	SD	0.51	0.5	<i>n. s.</i>
	N	188	387	
平均 成績	M	3.55 ^b	3.19 ^a	$F(1, 547)=23.90$
	SD	0.89	0.74	$p<.001$
	N	185	399	

注) アルファベットが異なる平均値の間には $p<.05$ で有意差がある。

表 4 . 授業曜日による授業評価と平均成績の平均値と標準偏差

	月	火	水	木	金	F値	
授業 評価	M	4.08	4.14	3.97	4.06	4.13	$F(4, 542)=1.57$
	SD	0.5	0.48	0.55	0.47	0.53	<i>n. s.</i>
	N	130	166	83	123	71	
平均 成績	M	3.29	3.33	3.1	3.31	3.29	$F(4, 542)=1.30$
	SD	0.83	0.79	0.78	0.75	0.71	<i>n. s.</i>
	N	126	155	81	117	68	

に教職 (4.30)・保健体育 (4.22)・英語 (4.21), 第 3 に国際文化 (4.12), 観光産業 (4.03), 経営情報 (4.00) の専門科目の値が高かった。一方, 平均成績では, 第 1 に保健体育 (3.89) の値が最も高く, 第 2 に教職 (3.69)・第二外国語 (3.51)・国際文化 (3.40)・経営情報 (3.36) が, 第 3 に観光産業 (3.44)・情報科目 (3.29) が続いた。なお, 両方の指標で共通し, 総合 ($M=3.80, 2.52$) と教養基礎 ($M=3.82, 2.81$) の 2 科目が低い値を示していた。

授業科目の必修・選択指定による違いを検討したのが表 3 である。授業評価では有意な差は認められなかった ($F(1, 547)=.44, n. s.$), 平均成績では必修指定の科目 ($M=3.55$) の方が選択指定の科目 (3.19) よりも有意に高かった ($F(1, 547)=23.90, p<.001$)。

曜日を独立変数とした分散分析結果は表 4 に掲載した。授業評価 ($F(4, 542)=1.57, n. s.$) ならび平均成績 ($F(4, 542)=1.30, n. s.$) とも, 有意な主効果は認められ

なかった。

科目の設定時限の結果は, 表 5 で検討した。授業評価では 1 時限から 6 時限の間に有意な差は認められなかった ($F(5, 541)=1.08, n. s.$)。しかし, 平均成績で主効果が有意となり ($F(5, 541)=9.97, p<.001$), 下位検定の結果, 5 時限 ($M=3.76$) と 6 時限 (3.83) に開講されている科目は, それ以外の科目 (1 時限から順に 3.13, 3.12, 3.20, 3.25) より平均成績の値が有意に高かった。

授業評価, あるいは平均成績の値について, 29 人以下の少人数授業と 30 人以上の大人数授業の間で差が存在するかを検討した (表 6)。その結果, 授業評価 ($F(1, 545)=54.95, p<.001$) と平均成績 ($F(1, 545)=111.99, p<.001$) ともに受講生数の大きさの主効果が認められ, 少人数授業 (授業評価 4.24, 平均成績 3.60) の方が大人数授業 (授業評価 3.93, 平均成績 2.96) よりも値が有意に高かった。

授業の対象年次間でその得点に違いがあるかを検討した (表 7)。その結果, 授業評価に有意な主効果が認められ

表5. 時限による授業評価と平均成績の平均値と標準偏差

		1時限	2時限	3時限	4時限	5時限	6時限	F値
授業評価	M	4.1	4.04	4.02	4.12	4.14	4.25	$F(5, 541)=1.08$
	SD	0.5	0.52	0.51	0.46	0.5	0.6	<i>n. s.</i>
	N	106	125	123	121	88	10	
平均成績	M	3.13 ^a	3.12 ^a	3.2 ^a	3.25 ^a	3.76 ^b	3.83 ^b	$F(5, 541)=9.97$
	SD	0.77	0.75	0.75	0.73	0.75	0.9	$p<.001$
	N	100	118	119	118	82	10	

注) アルファベットが異なる平均値の間には $p<.05$ で有意差がある。

表6. クラスサイズによる授業評価と平均成績の平均値と標準偏差

		少人数	大人数	F値
授業評価	M	4.24 ^b	3.93 ^a	$F(1, 545)=54.95$
	SD	0.51	0.44	$p<.001$
	N	287	286	
平均成績	M	3.6 ^b	2.96 ^a	$F(1, 545)=111.99$
	SD	0.8	0.61	$p<.001$
	N	273	274	

注) アルファベットが異なる平均値の間には $p<.05$ で有意差がある。

表7. 対象年次による授業評価と平均成績の平均値と標準偏差

		1年	2年	3年	4年	F値
授業評価	M	4.03 ^a	4.04 ^a	4.14 ^a	4.38 ^b	$F(3, 545)=5.05$
	SD	0.48	0.49	0.51	0.51	$p<.01$
	N	238	154	146	37	
平均成績	M	3.1 ^a	3.12 ^a	3.55 ^b	4.15 ^c	$F(3, 545)=34.23$
	SD	0.78	0.66	0.77	0.8	$p<.001$
	N	236	155	149	44	

注) アルファベットが異なる平均値の間には $p<.05$ で有意差がある。

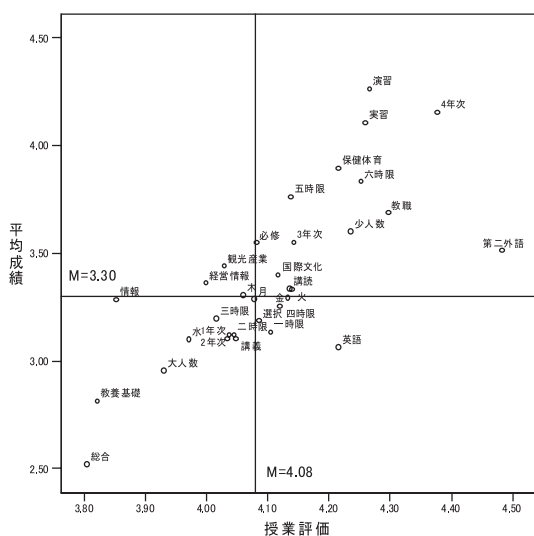


図3. 授業評価と平均成績の二次元上に布置した授業特性

($F(3,545)=5.05, p<.001$), 4年次対象科目 ($M=4.38$) が他年次対象科目 (1年次から順に4.03, 4.04, 4.14) よりも授業評価が有意に高かった。また、平均成績でも有意な主効果が認められ ($F(3,545)=34.23, p<.001$), 1年次 (3.10)・2年次対象科目 (3.12)よりは3年次対象科目 (3.55), さらには3年次対象科目よりも4年次対象科目 (4.15) において平均成績の値が高かった。

以上、7つの授業特性に基づく34の下位特性の平均値を算出した (表1~7を参照)。そこで、横軸を授業評価、縦軸を平均成績とした2次元上にその34の下位カテゴリーを布置した (図3)。参照線は図2-2と同様に、授業評価と平均成績の平均値 ($M=4.08, 3.30$) を用いた。その結果、相対的に授業評価が高く平均成績も高い、第1象限には、演習、実習、保健体育、5時限、6時限、教職、第二外国語などの科目群が布置された。一方、平均値から相対的に授業評価が低く、平均成績も低い第3象限には、総合、教養基礎、大人数などの科目群が布置された。授業評価は高くないが、平均成績は高い第2象限

表 8-1. 数量化 (授業評価のカテゴリカル回帰分析用)

	カテゴリ	数量化		カテゴリ	数量化
授業評価	2.38 - 3.27	-1.94	受講者数(人)	1 - 7	-1.36
	3.28 - 3.64	-1.3		8 - 27	-0.66
	3.65 - 3.94	-0.66		28 - 46	0.05
	3.94 - 4.22	-0.01		47 - 64	0.75
	4.23 - 4.52	0.63		66 - 90	1.45
	4.52 - 4.89	1.27		91 - 204	2.16
	5	1.91	曜日	月	-1.38
授業形態	講義	-0.42		火	-0.32
	講読	0.73		水	0.08
	演習	2.48		木	0.94
	実習	4.29		金	1.65
科目区分	英語	-1.3	時限	一時限	0.01
	第二外語	-2.23		二時限	1.05
	教養基礎	0.52		三時限	0.11
	情報	-0.09		四時限	0.34
	保健体育	-0.71		五時限	-2.19
	総合	0.79		六時限	0.53
	国際文化	0.5	受講年次	1年次	-1.14
	経営情報	0.71		2年次	0.63
	観光産業	0.66		3年次	0.63
	教職	-1.16		4年次	2.07
必修/選択	必修	-1.45			
	選択	0.69			

表 8-2. 授業評価に及ぼす授業特性の影響

	β	標準誤差	df	F	p
授業形態	.234	.057	3	16.736	.000
受講者数(人)	-.167	.046	1	13.257	.000
科目区分	-.416	.048	9	76.253	.000
必修/選択	.118	.049	1	5.766	.017
曜日	.007	.041	4	2.992	.018
時限	.087	.043	5	4.183	.001
受講年次	.120	.049	2	5.986	.003

($R^2=.236, p<.001$)

には、特にどの科目群も布置されなかった。ただし、第2象限と第3象限の境界線に情報科目の存在が認められた。最後に、授業評価は相対的に高いが、平均成績が低い第4象限に認められたのは英語科目であった。

なお、7種類の授業特性は、実は相互に関連し合いながら、授業評価や成績に影響を及ぼしていると考えられる。例えば、演習科目の多くは少人数の受講生で、かつ5時限あるいは6時限に開講されていることが多い。また、総合科目や教養基礎科目は大人数で1年次対象、あるいは1～4時限目に多くが開講されている。そのため、これらの授業特性間の相互関連を考慮し、各授業特性が持つ授業評価と成績に対する独自の影響度を評価する必

要がある。そのため、本研究ではカテゴリカル回帰分析を用いた。目的変数には授業評価あるいは平均成績を用い、説明変数には7つの授業特性を設定した。なお、クラスサイズはカテゴリ変数としてではなく、実際の受講生数を分析に用いた。

授業評価を目的変数としたカテゴリカル回帰分析において、各変数を表8-1の通り数量化した。授業評価を7つの授業特性が説明する程度の指標である重相関係数 R^2 は.236 ($p<.001$)であり、7つの授業特性が授業評価を有意に説明することが明らかとなった。各授業特性の標準偏回帰係数は表8-2の通りである。第1に、科目区分が授業評価に最も大きな影響を与えていた。数量化の絶

表9-1. 数量化 (平均成績のカテゴリカル回帰分析用)

	カテゴリ	数量化		カテゴリ	数量化		
平均成績	1.22 - 2.00	-1.89	受講者数(人)	1-7	-1.36		
	2.02 - 2.59	-1.24		8-27	-0.66		
	2.60 - 3.07	-0.58		28 - 46	0.05		
	3.08 - 3.53	0.07		47 - 64	0.75		
	3.53 - 4.00	0.73		66 - 90	1.45		
	4.01 - 4.60	1.38		91 - 204	2.16		
	4.62 - 5.00	2.04					
授業形態	講義	-0.4	曜日	月	-1.66		
	講読	-0.05		火	0.27		
	演習	2.6		水	0.59		
	実習	2.42		木	1.16		
				金	-0.15		
科目区分	英語	0.37	時限	一時限	-1.63		
	第二外国語	-2.47		二時限	0.33		
	教養基礎	0.17		三時限	0.68		
	情報	-1.1		四時限	-0.23		
	保健体育	-1.03		五時限	0.3		
	総合	1.98		六時限	4.42		
	国際文化	0.45	受講年次	1年次	-1.17		
	経営情報	0.16		2年次	0.46		
	観光産業	0.58		3年次	1.08		
	教職	-1.32		4年次	1.2		
	必修/選択	必修		-1.45			
		選択		0.69			

表9-2. 平均成績に及ぼす授業特性の影響

	β	標準誤差	df	F	p
授業形態	0.38	.048	3	63.318	.000
受講者数(人)	-.219	.040	1	29.519	.000
科目区分	-.230	.038	9	36.064	.000
必修/選択	-.055	.046	1	1.441	.231
曜日	-.045	.035	4	1.654	.159
時限	.104	.035	5	9.108	.000
受講年次	.169	.043	3	15.320	.000

($R^2=.401$, $p<.001$)

対値の大きい第二外国語、教職、英語が授業評価を高める方向で作用していることが分かる。2番目に大きな影響度を示したのは、授業形態であった。特に数量化において、絶対値の大きい実習が授業評価を高めることに貢献していた。3番目に影響度の強かった変数は、受講生数であった。受講生数が多くなるほど、授業評価が低くなることが明らかになった。

同様に、平均成績を目的変数とした場合のカテゴリカル回帰分析における各変数の数量化は、表9-1に示した通りである。重相関係数 R^2 は.401 ($p<.001$)であり、各授業特性の標準偏回帰係数は表9-2の通りであった。1番目に影響力が強かったのは授業形態であり、数量化における演習

と実習の絶対値の大きさから、この2つの科目が平均成績を高める方向で作用していることが明らかとなった。2番目に影響力が高かった変数は科目区分であり、数量化の絶対値の大きさから、第二外国語と教職の科目が平均成績を高める方向で作用し、総合科目が低める方向で作用していることが明らかとなった。3番目に影響力が高かった変数は受講生数であり、受講生数が多くなるほど平均成績が低くなっていた。一方で、必修・選択指定と曜日は影響力が有意でないことが明らかとなった。

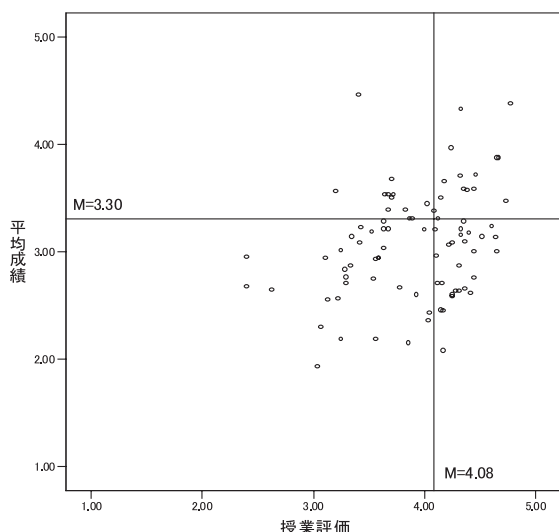


図4. ある科目区分の授業形態「講義」における図示例

考察

本研究では、大学全体のカリキュラムを構成する科目を、それぞれの受講生による授業評価および教員による成績評価の2次元上に布置することで、従来よりも利用しやすく、適切な授業改善の理解を促進するための図式を提案することを目的とした。提案した図式に基づき、授業評価および成績の両データが利用可能だった549科目を2次元上に布置した結果、大学カリキュラムの中で相対的に授業改善を要する科目（第3象限のR低G低に布置）と相対的に授業改善を要しない科目（第1象限のR高G高に布置）をリストアップすることができた。R高G高の領域には160科目、R高G低の領域には133科目、R低G高の領域には82科目、R低G低の領域には174科目が布置された。そして、授業評価と平均成績との間には弱い相関が認められたものの、授業評価の高い授業が成績（学習成果）の良好な授業であり、授業評価の低い授業が成績（学習成績）が良好でない授業であると単純に断定できないことが明らかとなった。

授業評価と成績の2次元布置によるカリキュラム把握は、従来の授業評価データに依存したものに比べ、各授業を4つの位相に位置づけるため、授業改善の方策を考える良いツールとなるであろう。例えば、図4は某科目区分の「講義」に属する87科目を2次元上に布置したものである。この図を授業改善の資料として用いれば、カリキュラム編成責任者は当該科目のカリキュラム編成において、授業評価の低い科目ないし学習成績の低い科目の共通点を見出し、単独の授業改善を促すことはもちろんのこと、前提科目の設定や導入科目の新設など、全体的な視野に立って編成計画の立案が可能となるであろう。さらには、教員個人においても、カリキュラム全体の中

で自分の相対的位置づけを確認することは、単純に授業評価データをフィードバックされるよりも授業改善の意識を高めるであろう。

なお、この2次元布置を利用する際に注意すべき点は、シラバス（授業計画書）において、全体の教育目標はもちろんのこと、各授業の教育目標と成績評価が明確化されている必要がある。小笠原ら（2002）も述べている通り、授業評価の導入は、シラバスの質が保証されていなければ授業改善効果を発揮しない。もし、大学の教育目標が不明確であったり、教育目標を実現するための体系的なカリキュラム編成が不十分であったり、さらには、教員自らが大学側の承認を受けた授業計画書を遵守しなければ、今回の2次元布置の資料も、期待しなかった方向での大学カリキュラムの改悪を生むであろう。つまり、教員が望ましい授業（R高・G高）を実現するために、意識的、あるいは無意識的にかかわらず、大学の教育目標とは無関係に授業目標を低水準に設定したり、到達度の低い学生に対しても良い成績を与えたりするという危険性が存在する。

本研究のもう1つの目的は、授業固有の特性（外部要因）が成績と授業評価に及ぼす影響度を検討することで、授業評価と成績に基づき、教員個人の教育活動を評価する際の問題点について明らかにすることであった。そのため、授業評価および平均成績を従属変数、それぞれの授業特性を独立変数とした一要因分散分析を繰り返し行い、さらに、授業特性ごとの平均値を2次元上に布置し、最後にカテゴリカル回帰分析によって授業特性の独立した影響度を評価しようと試みた。一要因分散分析の結果、授業評価に影響する授業特性として、授業形態、科目区分、クラスサイズ、対象年次の4つが見出され、平均成績に影響する授業特性として、授業形態、科目区分、必修・選択指定、授業時限、クラスサイズ、対象年次が見出された。そして、授業評価と平均成績の2次元上にそれら34の授業特性を布置した結果、授業特性によって、授業評価および平均成績の得点に大きな違いが現れることが明らかとなった。さらに、カテゴリカル回帰分析の結果、授業評価に対しては、科目区分、授業形態、受講者数の順に影響力が強く、平均成績に対しては、授業形態、科目区分、受講者数の順に影響力が強いことが明らかとなった。

この結果は、実に興味深い結果を示唆している。つまり、授業評価を教員の教育業績の資料として利用する場合には、単純にその数値で教員間を比較するのではなく、授業形態、科目区分、受講者数を考慮し、授業評価得点を補正して活用することの必要性を示唆している。例えば、授業評価において、第二外国語・英語・教職科目、演習・実習科目、少人数クラスを担当している教員は、相対的に高い授業評価を得やすい授業を担当していると

考えられる。一方で、総合・教養基礎科目、講義・講読科目、そして大人数クラスを担当している教員は相対的に授業評価が低くなりやすい。この授業特性は教員個人が改善しようと思っても容易に改善できるものではないため、単純に授業評価の結果に基づいて、教員個人の教育業績を査定することは困難だといえる。この結果は、様々な授業の特性が授業評価に影響を及ぼすとする先行研究の知見とも一致する。さらに、本研究はその先行研究の結果を拡張し、同様の授業特性が平均成績にも影響することを明らかにした。この点は、本研究のオリジナルな発見である。実際の今後の課題としては、高い授業評価と高い学習成果が得られやすい科目とそうでない科目が存在することを考慮した上で、カリキュラム編成、授業改善および教員の教育業績評価のあり方を検討する必要があるであろう。例えば、南(2004)は、教育業績評価を行うにあたり、5段階評定の授業評価得点を受講生一人につき0.002加算する(100人につき0.2ポイント加算する)必要性を報告している。もちろん、各大学によって授業評価の質問項目やその他の状況に違いがあるため、補正数値については個別に検討する必要があるであろう。

最後に、本研究の問題と課題を3つ挙げて終わりたい。第1に、本研究で用いた授業評価ならびに成績評価の信頼性と妥当性の問題がある。本研究の授業評価および成績の指標は、非常に簡便なものであった。授業評価は1項目(「この授業の総合評価をして下さい」)に対する5段階評定)で測定され、成績評価は教員が行った5段階評価に基づいていた。それらの得点の信頼性と妥当性について疑問視する読者も多いであろう。例えば、授業評価の調査は期末試験と同時に実施されなかった。そのため、本来授業を適正に評価することが難しい、出席回数少ない受講生の回答が授業評価に影響した可能性が残る。なお、授業評価アンケートで収集された欠席の自己報告(5段階評価で、「5 一度も休んだことはない」~「14回以上」)の平均値と授業評価平均との相関係数を算出したところ有意ではなかった($r=.036$, $n.s.$, $n=575$)。つまり、本研究のデータは、受講生の欠席度の影響をそれほど受けていないと推測できる。しかしながら、今後、授業評価得点の信頼性・妥当性を高めるためにも不適切な受講生による授業評価の回答の除去が重要であろう。また、実際の授業評価の質問項目は総合評価だけでなく、教員特性、授業内容、授業システムなどを評価するための12項目を同時に測定していたが、本研究の分析では利用しなかった。今後、この下位項目を利用した詳細な分析を進める必要があるだろう。

他方、成績に関しては、当初の教育目標を本当に達成したか否かは、教員の判断に拠るところが大きい。最終的な成績評価を算定する方法にも複数あり、出席、授業への積極的参加度、課題発表、レポート、小テストなど

多様化しており、第三者が当初の学習成果の達成度を成績のみから判断することは難しいといえる。教育の質の保証を考えた場合、例えば、授業計画書への教育目標と成績評価基準の具体的な明記、受講生や自己点検・評価委員会からの公開請求に応える成績評価資料(例、出席簿、提出課題や期末試験等の採点結果)の作成、成績評価を保証する第三者審査等、成績評価の信頼性と妥当性を全学的に保証するシステムの必要性がでてくるであろう。

第2に、受講生や教員、授業の内部要因によってこの2次元布置が影響を受ける可能性が考えられる。受講生が自発的に授業計画書を読み、受講しているならば、授業評価は授業計画書を基準に左右され、成績が高いか低いかに影響を受けないであろう。一方で、受講生が単に単位取得を目的に受講している場合は、授業評価は授業計画書や授業内容によって影響を受けるよりも、単位取得の容易さによって影響を受けるかもしれない。さらには、受講生側の要因は、教員の熱意、対話、教材の利用、授業計画書の提示などといった教員側の要因と複雑に相互作用しながら、授業評価および成績に影響を及ぼすと考えられる。今後、外部要因だけでなく、受講生の要因と教員の要因を積極的に分析に組み込む必要があるであろう。

なお、授業改善において重要となってくる指標は、改善の余地が残されているにもかかわらず、総合的な授業評価ないし成績と相関が高いものと考えられる。例えば、本研究で扱った授業の特性のように、教員の年齢、職階などは改善の余地がほとんどないため、授業改善の際に重要とならない指標である。一方、受講生の受講態度、教員の熱意や対話、授業システム(教員の遅刻、授業の休講、補講)が授業評価や成績と強い相関関係が認められるにもかかわらず、それらが低水準にある科目は、授業改善の緊急性が高いと判断される。今後、授業評価と成績との間に相関関係が認められる指標に基づき、2次元布置をさらに精緻化する必要がある。

本研究が提案した2次元布置によるカリキュラム理解は、あくまで従来の授業評価研究に対する1つの改善策としての提案であった。しかし、従来の研究と同様、教育評価の精度を高めることだけでは本末転倒である。最後の課題として、2次元布置によるカリキュラム評価が、具体的な授業改善の方策の提案や大学全体の授業改革の意識高揚につながる仕組みづくりが残されているであろう。そのような目的を達成するには、例えば、カリキュラム全体での学習指導形態(一斉学習、グループ学習、個別学習)による影響というマクロ的視点はもちろん、複数の同一科目における実験授業を詳細に検討することも必要となってくるであろう。また、本研究では考慮しなかった、絶対評価、相対評価、個人内評価といった成

績評価の種類による差異を検討することも必要となってくるであろう。望ましい授業とは何かを考えた場合、授業後の到達度だけではなく、授業前後の学力向上の側面をも考慮する必要がある。さらには、教員や学生自身が授業評価システムを十分に理解した上で利用していくためのマニュアル作りや、授業評価や成績評価の信頼性や妥当性に対する様々な質問に答えられるだけの研究成果の蓄積が今後の課題となってくるであろう。

付 記

本研究は、平成15年度名城大学総合研究所（一般研究）から研究助成を受けて行われた。

引用文献

- 牧野幸志 2001a 学生による授業評価と自己評価、成績、及び学生の満足感との関係 教養選択科目「社会心理学」の場合 高松大学紀要, 35, 1-16.
- 牧野幸志 2001b 学生による授業評価と自己評価、成績、及び学生の満足感との関係 専門必須科目「人間関係論」の場合 高松大学紀要, 35, 17-31.
- 牧野幸志 2001c 学生による授業評価の規定因の検討 (1) 多変量解析を用いた因果モデルの検討 高松大学紀要, 36, 55-66.
- 牧野幸志 2002 学生による授業評価、満足感と成績との関係 成績の悪い学生は本当に授業を酷評するのか? 高松大学紀要, 38, 35-47.
- 牧野幸志 2005 学生による授業評価と出席率との関係 (1) 授業に出ていない学生は授業を悪く評価するのか? 経営情報研究 (摂南大学経営情報学部論集), 13, 1-14.
- 松田文子・三宅幹子・谷村 亮・小嶋佳子 1999 学生による授業評価と自己評価、授業選択態度、及び成績の関係 教職必修科目「生徒指導論」の場合 広島大学教育学部紀要 (第一部) 心理学, 48, 121-130.
- 南 学 2004 学生による授業評価におけるクラスサイズの効果 松山大学論集, 16, 57-75.
- 村田 治 2004 関西学院大学におけるファカルティディベロップメントへの取り組み Web 上でのシラバスと授業評価システムの連携例 大学教育と情報, 12 (3), 5-7.
- 西山裕美子 2005 共通教育成績評価の実態調査 高等教育システムセンター紀要, 1, 65-76.
- 小笠原昭彦・小玉香津子・生田克夫・大平政子・北川真理子・勝又正直 2002 シラバスの向上と学生の授業評価による教育改善について 名古屋市立大学看護学部におけるファカルティ・ディベロップメントの試み 名古屋市立大学看護学部紀要, 2, 129-137.
- 櫻井公人 1998 教育ノウハウ・データベース構築の意義 FD としての授業評価、シラバスとの関連で 大学教育研究所年報 (1), 3-7.
- 浦上昌則・林 雅代・石田裕久 1999 受講動機別にみた授業評価と満足度 アカデミア (人文・社会科学編), 70, 515-540.
- 渡辺勇一 2001 学生による授業評価をどう見るか 生物科学, 52 (4), 209-216.
- 安岡高志・吉川政夫・高野二郎・峯崎俊哉・成嶋 弘・光澤舜明・道下忠行・香取草之助 1989 学生による講義評価 学生の質と講義評価の関係について 東海大学紀要 (教育研究所教育工学部門), 2, 31-34.
- 八ッ橋武明 2000 授業評価における授業満足度の構成要因情報研究 (文教大学情報学部編), 25, 163-176.