

学生による授業評価に影響を及ぼす諸要因について Some Factors that Affect the Course Evaluations by Students.

藤田 勉^{*1 §}, 川島 眞^{*2}

Tsutomu FUJITA, Makoto KAWASHIMA

Abstract: Through six studies, we investigated some factors that may affect the course evaluations by students. In Study 1, statistically significant relationships have been detected between the course evaluations and students' actual achievement on examination, students' self-ratings of attendance, self-ratings of learning attitude and self-ratings of expected performance on examination. In Study 2 and Study 3, we tried to find possible correlations between student evaluations and their psychological traits with an Egogram. Total scores of Egogram (the sum total of CP, NP, A, FC, AC scores) correlated with student evaluations in three out of four courses, although scores at each category of Egogram didn't have consistency across courses. In Study 4, we investigated the relationship between student evaluations and the overall impression that the lecturer gave to students. The results showed a .39 correlation coefficient ($p < .001$) between student evaluations and overall impression of lecturer. The relationship between student evaluations and student anonymity was examined in Study 5, but student anonymity didn't significantly influence student evaluations. In Study 6, we asked students to evaluate the same course twice, and students evaluated the course at the end of course (1st rating) and again sixteen months or seven months later (2nd rating). The retrospective ratings tended to be lower scores than the 1st rating in most questions. The results of the present studies suggested that some factors could be biasing variables on student evaluations. It seemed that we should know biasing variables in order to utilize student evaluations effectively.

Keywords: course evaluations, factors, correlation, Egogram, students

はじめに

2013年11月7日に文部科学省から発表された「大学における教育内容等の改革状況等について」によれば、2011年度全ての学部もしくは全ての研究科で学生による授業評価を実施した大学数は708大学で、これは国公立大学全体の約93%に相当する。また、授業評価に関するアンケート結果を組織的に検討し、授業内容等に反映する機会を設けているのは368大学(全体の48%)であった(文部科学省高等教育局大学振興課大学改革推進室, 2013)。

学生による授業評価は最初にアメリカの大学で導入された。アメリカにおける学生による授業評価導入の歴史についてはAlgozzine, Beattie, Bray, Flowers, Gretes, Howley, Mohanty, & Spooner (2004), Centra (1993), 大山 (2007), Marsh (1987), Wachtel (1998)などに詳しい。それらの文献によ

ると、1920年代の中頃にはハーバード大学(Harvard University), ワシントン大学(University of Washington), パデュー大学(Purdue University), テキサス大学(University of Texas)などで学生による授業評価が導入され(Marsh, 1987), 中でもインディアナ州にあるパデュー大学では, Remmersらが中心となり独自の評価スケール(the Purdue Rating Scale for Instructors)を開発し, 実証的な分析をしていたようである(Algozzine et al., 2004; Centra, 1993; Marsh, 1987; Wachtel, 1998)。その後、1960年代には多くの大学で学生による授業評価が実施されるようになり、1970年代に入ると学生による授業評価の妥当性や有効性に関する研究が盛んに行われるようになった。それ以降は、先行研究の結果を確認・発展させる研究、あるいは関連する研究結果を統合するような研究が数多く報告されている。アメリカにおいて早くから学生による授業評価が一般的になった背景には、「学生消費者主義(student

*1 長野県短期大学幼児教育学科 *2 尚美学園大学芸術情報学部音楽表現学科

§ 連絡先 〒380-8525 長野県長野市三輪8-49-7 TEL 026-324-1221 FAX 026-235-0026

consumerism)」の考え方があるとされている（喜多村, 1990, 1992）。つまり、学生を大学からのサービスを受ける消費者としてとらえ、授業料を支払っている学生には“質の高い教育を受ける権利”，授業料を受け取っている大学側には“質の高い教育を提供する義務”があり，そのためには大学サービスの主たる“商品”である授業の質を査定する必要があるというわけである。

我が国における学生による授業評価の導入については、安岡（2005）にまとめられている。安岡によれば、学生による授業評価の導入は「日本では、1974年の国際基督教大学（ICU）が最初で、続く1984年に東海大学が実施しているが、この段階では、まだ一部の有志が行うにとどまっていた。やがて、1988年の国際基督教大学を皮切りとして、組織的に導入する大学が増え始め、1990年には多摩大学と慶應義塾大学 SFC、1993年の東海大学と続く」とある。また、井上（2010）は、「日本で初めて1つのキャンパスにおいて開講科目のすべてで、一斉に学生による授業評価調査が導入されたのは、慶應義塾大学の湘南藤沢キャンパス（SFC）であり、それは平成2（1990）年のことであった」としている。

このように、我が国では1990年代以降多くの大学が学生による授業評価を導入するようになったが、その背景には1991年の大学設置基準の大綱化、1999年の大学設置基準においてFD（ファカルティ・ディベロップメント）が努力義務になったこと、2004年の認証評価制度の創設、2007年の大学院設置基準においてFDが義務化されたこと、2008年度から学士課程でもFDが義務化されたことなどがあったと考えられる。

大学教員の多くは大学院等で“研究すること”の訓練を受けてきたが、“教育すること”は教えられてこなかった。そのため、効果的な授業を展開するためには、実際に教壇に立つようになってから自分自身でティーチング・スキルを磨く必要に迫られる。アメリカの大学では、若手研究者に対し、ティーチング・スキルを向上させるためのマニュアルを作成して配布しているところもあるが（谷口, 2005）、日本の大学においてはそのような教育能力を向上させるための組織的な教育はほとんど行われてこなかった。

こうした状況を受け、中央教育審議会は、2005年の答申「新時代の大学院教育」で、「確かな教育能力と研究能力を兼ね備えた大学教員の養成」を大学院に求められる人材養成機能の1つとして明示した（中央教育審議会, 2005）。その中で、大学院には

「これまで脆弱であった教育を担う者としての自覚や意識の涵養と学生に対する教育方法等の在り方を学ぶ教育を提供することが求められる」とし、例えば、ティーチングアシスタント（TA）等の活動を通じて、授業の実施方法や教材の作成に関する教育などを実施するといった方策を提案している。

このような動きもあり、これからは“教育すること”の訓練を受けてから大学教員になる研究者も増えることが期待されるが、そうした訓練を受けてこなかったこれまでの大学教員は何らかの方法で、自らの授業方法や授業内容の改善をしなければならない。そのための1つの手段になり得ると目されるのが学生による授業評価である。しかし、この学生による授業評価については、様々な問題点が指摘されている。例えば、「授業評価調査を実施はしているが、形式だけ・体裁だけであることも多い」（井上, 2010）、「各大学で授業評価を初めて導入する際には、多くの場合、授業の改善ということがまず標榜されるが、現実には大学評価のためにということが先にあって、授業改善は名目で終わっている大学も少なくないのが現状」（大塚, 2010）、「教員からも事務スタッフからも学生授業評価をやってよかったという声が聞こえてこない」（米谷, 2010）、「授業評価の結果はほとんどの大学が公表しているが、それは自己申告に基づくもので、その実態は大学によって雲泥の差がある」（小田, 2010）などである。

本学においては、これまでFD委員会（2004年度は大学評価委員会）を中心に2004年度、2005年度、2008年度、2009年度、2010年度、2011年度、2014年度に「授業改善アンケート」というかたちで学生による授業評価が実施された。藤田らは、それに先立つ1992年度から本学で開講された授業を対象に学生による授業評価を実施し、授業評価に及ぼす諸要因について検討してきた（藤田・川島, 1993；藤田・久東・川島, 1994；藤田・西村・清水・小林・久東・川島, 1995；藤田・久東・川島, 1995；藤田・久東・川島, 1996；藤田・久東・川島, 1997；川島・久東・藤田, 1994；川島・久東・藤田, 1995；川島・久東・藤田, 1996）。本稿では、藤田らの研究で検討された学生による授業評価に影響を及ぼすいくつかの要因を紹介する。本学で実施された学生による授業評価の特徴を知ること、学生による授業評価を有効に活用する上で参考になると思われる。

研究1

学生による授業評価に影響を与える変数としては

様々なものが考えられる。例えば、同じ授業を受講しても、その学生の理解度、出席状況、授業態度、単位のとりやすさに対する認識の違いなどによっても授業評価の結果は大きく変わってくるのが予想される。こうした変数のうち、阿部（1987）は試験の予想成績と実際の成績について検討しており、試験の成績が良いと予想した学生ほど授業評価が高く、実際の成績が良い学生ほど良い授業であったと評価する傾向があることをみいだした。また、竹綱（1992）は出席率と授業評価との関係を検討し、出席率が高い学生ほど良い授業評価をしやすいことを示した。

研究1では、そうした先行研究の結果をふまえ、授業評価に影響を及ぼす諸要因について検討した。具体的には、一般教育科目の「心理学」を受講した短期大学生を対象に授業評価に関するアンケート調査を行い、授業評価の結果と試験の得点、出席状況・授業態度・試験についての自己採点との関連を調べた。また、単位のとりやすさに関しても質問し、単位のとりやすさに対する認識の違いが授業評価の差となって表れるかどうかについても検討した。

方法

対象者 1993年度前期に筆者の1人（藤田）が担当する一般教育科目「心理学」を受講した女子短期大学生257名。

手続き 「心理学」の試験時に試験用紙とともに授業評価に関する質問項目が書かれた用紙（授業評価シート）を配付し調査を実施した。質問項目は、すでに学生による授業評価を実施している大学や先行研究で用いられている項目を参考にして作成した。授業評価を求める質問項目には、①授業はわかりやすかったか、②授業はまとまっていたか、③授業の進み具合はどうだったか（早すぎたり、遅すぎたりすることはなかったか）、④教員の熱意は感じられたか、⑤教員は授業の準備を十分にしていたか、⑥トピックの選び方は適切だったか、⑦板書は見やすかったか、⑧声の大きさはどうだったか（大きすぎたり、小さすぎたりすることはなかったか）、⑨話すスピードはどうだったか（早すぎたり、遅すぎたりすることはなかったか）、⑩教材（教科書や授業中に用いた資料など）はどうだったか、の10項目があり、それぞれの質問に対して0点（非常に悪い）～10点（非常に良い）で評定するよう求め、出席状況・授業態度・試験についても0～100点で自己採点してもらった。また、「心理学の成績のつけかたについてどのように思うか」、「今日のテスト

の難易度はどうだったか」の項目をもうけ、前者は0（とてもやさしそう）～100点（とても厳しそう）、後者は0点（とてもやさしい）～100点（とても難しい）の間で評定させた。

授業評価と実際の試験得点との関係を検討するため、授業評価シートへの記名を求めたが、授業評価シートを配付した際、各質問項目に対する回答は成績をすべてつけ終わってから集計することを学生に伝え、各質問項目に対してどのような答え方をしても成績にはまったく影響しないことを強調して説明した。なお、当日の試験は30問の穴埋め問題（配点は各問3点）からなり、授業評価シートのすべての質問項目に回答した者には試験の得点に10点を加点する旨学生に伝えた。

結果と考察

授業評価に関する質問10項目の評定値を合計したものを授業評価の総合点とした。総合点は最低で61点、最高で100点で、平均は92.20点（SD=8.08）であった。

Table 1

授業評価の総合点と各変数間の相関係数

試験の得点	.26**
出席状況の自己採点	.22**
授業態度の自己採点	.28**
試験の自己採点	.15*
成績のつけかた	-.01
試験の難易度	-.05

* $p<.05$, ** $p<.01$

授業評価の総合点と実際の試験の得点、および前述の各質問項目に対する評定値間の相関係数を算出した。結果をTable 1に示す。授業評価の総合点と有意な相関があったのは、試験の得点（ $p<.01$ ）、出席状況の自己採点（ $p<.01$ ）、授業態度の自己採点（ $p<.01$ ）、試験の自己採点（ $p<.05$ ）であった。これより、「試験の得点が高い学生ほど授業内容や授業形式を高く評価しやすい（良い授業評価をする学生ほど試験の得点が高い）傾向」が確認された。また、「出席状況・授業態度・試験の得点についての自己採点が高い学生ほど良い授業評価をしやすい（良い授業評価をする学生ほど出席状況・授業態度・試験の得点についての自己採点が高い）」という傾向も認められた。一方、成績のつけかた（単位のとりや

すさ)や当日の試験の難易度に関する認識の違いは授業評価の総合点と明白な関係をもたなかった。

本研究の結果からも推測されるように、学生による授業評価は必ずしも授業の内容や形式だけを正確に反映するものではなく、学生側の様々な変数によっても影響を受けることが考えられる。学生による授業評価を実施する際には、学生の“授業評価行動”に影響を及ぼす変数を明らかにし、その妥当性を詳細に検討する必要があると思われる。

研究2

学生による授業評価は必ずしも授業の内容や形式だけを正確に反映するものではなく、学生側の様々な変数によっても影響を受けることが考えられる。そうした変数の1つに学生の性格特性(行動傾向)がある。本研究では、学生に授業評価を求めると同時に、性格検査を実施し、授業評価と性格特性間の相関関係について検討した。

学生の性格特性を測る指標としてはエゴグラム(山本, 1990)を用いた。エゴグラムはアメリカの精神科医 Eric Berne が創始した交流分析(Transactional Analysis: TA)の理論に基づいて John M. Dusay が開発した心理テストで、すべての観察可能な行動を5つの自我状態(尺度)に分類し、それらの発生頻度をグラフ化したものである(新里, 1986)。5つの自我状態とは、CP (Critical Parent 父親的自我状態:「批判的, 厳格, 命令的」な行動傾向), NP (Nurturing Parent 母親的自我状態:「保護的, 養育的」な行動傾向), A (Adult 大人の自我状態:「冷静, 客観的, 合理的」な行動傾向), FC (Free Child 自由性:「天真爛漫, 自由気まま」な行動傾向), AC (Adapted Child 順応性:「我慢強い, 他人に追従的」な行動傾向)で、これら5つの自我状態のバランスにより個人の性格特性を測ることができる。

方法

対象者 研究1で対象となった女子短期大学生 257名。

手続き 研究1で得られた授業評価の結果と「心理学」の授業時間中に実施したエゴグラムの結果をデータとして用いた。エゴグラムは性格心理学の説明の中で、質問紙法による性格検査の例として実施した。その際、この性格検査の各尺度の意味やエゴグラム・パターンの見方などについては検査実施後に説明した。

結果と考察

対象となった女子短期大学生 257名について、授業評価の総合点とエゴグラムの5つの尺度間の相関係数を求めた。結果を Table 2 に示す。

エゴグラムの5つの尺度のうち、授業評価の総合点と有意な相関がみられたのは、NP ($p<.01$), A ($p<.05$), FC ($p<.05$) の3尺度であった。すなわち、同じ授業を受講していても「保護的, 養育的」, 「冷静, 客観的, 合理的」, 「天真爛漫, 自由気まま」といった性格特性を強くもつ学生ほど授業を高く評価しやすいという傾向が認められた。中でも NP は5尺度の中で最も高い相関を示しており、NP の値と授業評価の総合点が深く関連していることが示唆された。

エゴグラムの開発者である Dusay は、CP, NP, A, FC, AC の得点総和を心的エネルギーの総量としてとらえており、5つの自我状態が変化したとしても、その合計は常に一定であるとしている(デュセイ, 1980; 新里, 1986)。本研究でも、その5つの尺度の得点総和と授業評価の総合点の間の相関係数を求めたところ $r=.26$ となった ($p<.01$)。このことから「心的エネルギーの総量が高い学生ほど良い授業であったと評価しやすい (= 良い授業評価をする学生ほど心的エネルギーの総量が高い)」ことがわかる。

また、エゴグラムの5つの尺度値と試験得点の間でも相関係数を算出したが、有意な相関がみられたのは CP ($r=.12, p<.05$) においてのみであった。

本研究では、授業評価の総合点と性格特性間の相関関係について検討したが、エゴグラムの5つの尺度のうち3つの尺度 (NP, A, FC) において授業評価との間に有意な相関がみられた。こうした傾向が「心理学」の授業だけに限らず他の授業科目につい

Table 2
授業評価の総合点とエゴグラム得点間の相関係数

CP	NP	A	FC	AC	5尺度の得点総和
.11	.19***	.14**	.15**	.03	.26***

** $p<.05$, *** $p<.01$

でも一貫してみられるのか、つまり、NP, A, FCが高い学生は他の科目の授業を評価した際にも高い評価をしやすいのか否かについて確認する必要があると思われる。

研究3

研究2では、エゴグラムでとらえられた学生の性格特性（行動傾向）と授業評価の総合点との関連性について検討した。その結果、エゴグラムの5つの尺度のうち3つの尺度（NP, A, FC）において授業評価との間に有意な相関がみられた。すなわち、同じ授業を受講しても、「保護的、養育的」、「冷静、客観的、合理的」、「天真爛漫、自由気まま」といった特性を強くもつ学生ほど授業内容や授業形式を高く評価する傾向が認められた。また、5つの尺度の得点総和と授業評価の総合点間の相関係数も有意であった。研究3では、こうした傾向が「心理学」以外の他の科目についても一貫してみられるか否かについて検討した。

方法

対象者 1993年度前期に一般教育科目「哲学」、「教育学」、「情報科学」を受講した女子短期大学生。このうちエゴグラム・データが得られている学生（「哲学」履修者114名中108名、「教育学」履修者95名中93名、「情報科学」履修者208名中203名）を対象に分析を行った。

手続き 各科目の試験時に試験用紙とともに授業評価シートを配付し、授業に対する評価を求めた。授業評価シートは、研究1、研究2で用いたものと同じものである。授業評価を求める際、学生に与える教示を科目間で統一するため、「哲学」、「教育学」、「情報科学」の担当教員にはあらかじめ教示内容が書かれた用紙を渡し、それを読み上げてもらった。エゴグラム・チェックは上記3科目と同時期に開講された「心理学」の授業中に実施した。

結果と考察

各科目に対する授業評価の総合点とエゴグラム得点間の相関係数を算出した。結果をTable 3に示す。Table 3をみてもわかるように、「哲学」では授業評価の総合点と5尺度の得点総和間、「教育学」では授業評価の総合点とAC間、「情報科学」では授業評価の総合点とCPおよび5尺度の得点総和間で有意な相関が認められた。研究2における「心理学」に対する授業評価では、授業評価の総合点とNP, A, FCの3つの尺度間で有意な相関がみられたが、本研究で対象となった3科目についてはそのような傾向は認められなかった。しかし、研究2においてみられた授業評価の総合点と5尺度の得点総和間の有意な相関関係は本研究の「哲学」および「情報科学」においても確認され、「心理学」とあわせて4科目中3科目で「エゴグラムの5尺度の得点総和が高い学生ほど良い授業であったと評価しやすい（＝良い授業評価をする学生ほど5尺度の得点総和が高い）」という傾向がみられたことになる。エゴグラムの開発者であるDusayは、5尺度の得点総和を心的エネルギーの総量としてとらえ、5尺度の総和が大きいくほど自我のエネルギーが強いとしているが、学生の授業評価行動にはこの心的エネルギーの総量が関与していると考えられる。この心的エネルギーの総量が意味するものを具体的に定義することは困難であるが、心的エネルギーの総量が学生個人の特徴を表すものであることは間違いない。4科目中3科目において、授業評価の総合点と心的エネルギーの総量の間には正の相関がみられたということは、学生の個人的な特徴が授業評価と関係しており、学生による授業評価が必ずしもその授業の内容や形式だけで決まるものではないことを示している。学生による授業評価は本来授業を評価するものであるから、授業内容や授業形式が独立変数であるべきだと思われるが、こうしたデータから、学生側のなんらかの要因が授業評価の内容に影響を及ぼしていることがわかる。

Table 3
授業評価の総合点とエゴグラム得点間の相関係数

	CP	NP	A	FC	AC	5尺度の得点総和
哲学	.07	.14	.18	.09	.03	.21*
教育学	.05	-.01	.08	-.06	.33***	.16
情報科学	.15**	.13	.07	.06	.06	.20***

* $p < .05$, ** $p < .01$

研究 4

研究 2 と研究 3 では、授業内容や授業形式以外の要因が学生による授業評価の結果に影響する可能性が示唆された。研究 4 では、学生が教員に対してもつ全体的な印象（好感度）に着目し、授業評価との関連について検討した。

方 法

対象者 1994 年度前期に一般教育科目「心理学」を受講した女子短期大学生 253 名。

手続き 研究 1～研究 3 で用いた授業評価シートの質問項目に加え、新たに学生が教員個人に対してもつ全体的な印象を好感度という尺度でとらえ、0 点（非常に悪い）～100 点（非常に良い）で評定するように求めた。「試験の難易度」、「成績のつけかた」、「出席状況・授業態度・試験についての自己採点」についても 0～100 点の間で評定してもらった点、授業評価を実施する際に留意した点などは、これまでの研究と同様である。

結果と考察

授業評価の総合点と試験の得点および各質問に対する評定値間の相関係数を求めた。結果を Table 4 に示す。

Table 4
授業評価の総合点と各変数間の相関係数

試験の得点	.15**
出席状況の自己採点	.25****
授業態度の自己採点	.26****
試験の自己採点	.18**
教員に対する好感度	.39****
成績のつけかた	-.03
試験の難易度	.07

** $p < .05$, *** $p < .01$, **** $p < .001$

研究 1 の結果と同様、本研究でも授業評価の総合点と試験の得点 ($p < .05$)、出席状況の自己採点 ($p < .001$)、授業態度の自己採点 ($p < .001$)、試験の自己採点 ($p < .01$) の間に有意な相関がみられた。しかし、授業評価の総合点と最も強い相関関係を示したのは新たに追加された変数である教員に対する好感度で ($p < .001$)、「教員に対する好感度が高い

（低い）学生ほど授業評価の総合点も高い（低い）」という傾向が認められた。また、教員に対する好感度は出席状況の自己採点および授業態度の自己採点とも強い相関がみられ（それぞれの相関係数は、 $r = .20$, $r = .25$ ）、「教員に対する好感度が高い（低い）学生ほど出席状況・授業態度の自己採点が高い（低い）」ことがわかった。

学生に授業評価を求めることは授業内容や授業形式の改善に役立つと思われるが、その反面、教員が学生からの評価を気にするあまり単位の認定基準を甘くしたり、人気とりに走ったりするなどの危険性をはらんでいる点が指摘されている (Marsh, 1987; Naftulin, Ware, & Donnelly, 1973)。本研究では、授業評価と教員に対する好感度の間に強い正の相関がみられたが、この結果は、教員に対する好感度を高めることがより良い授業評価につながる可能性を示唆するものである。好感度を高めるために授業の質的改善に努めるか、あるいはそれ以外の努力をするかは、教員自身の選択に依存している。

研究 5

研究 1～研究 4 では、授業評価の総合点と実際の試験得点間の相関関係の有無を検討するため、学生に授業評価シートへの記名を求めた。つまり、学生は匿名性が保持されていない状況で授業評価を行ったわけで、そのことが授業評価をする上で何らかのバイアスとなっていた可能性もある。匿名性が保持されている状況とそうでない状況での人間行動には少なからず差があることは社会心理学の分野の様々な研究からも明らかであり (Milgram, 1965; Zimbardo, 1970; Zimbardo, Haney, Banks, & Jaffe, 1977)、特に研究 1～研究 4 のように最終的な成績がつけられていない（単位が認定されていない）時点で授業評価を求める場合には、匿名性が保持されているか否かが授業評価の内容に大きく関わってくることも推測される。本研究では、このことを検討するため、研究 4 の授業評価（記名評価）の直後に再度学生に無記名で当該授業の評価をしてもらい（無記名評価）、両者の間に何らかの差があるか否かについて検討した。

方 法

対象者 1994 年度前期に一般教育科目「心理学」を受講した女子短期大学生 253 名。

手続き 研究 4 で学生に依頼した授業評価シートを回収した直後、再度授業評価シートを配付し、同

Table 5
記名評価と無記名評価の差

	記名評価	無記名評価	差 (記名 - 無記名)
授業評価の総合点	92.26	91.87	0.39
出席状況の自己採点	92.26	90.75	1.51
授業態度の自己採点	83.46	82.17	1.29
試験の自己採点	58.33	58.54	-0.21
教員に対する好感度	92.45	92.13	0.32
成績のつけかた	53.06	52.10	0.96
試験の難易度	67.06	66.33	0.73

じ質問項目に対して無記名で回答してもらった。提出は自由としたため回答が得られたのは 253 名中 240 名であった (回収率は 94.9%)。

結果と考察

研究 4 で得られた記名評価の結果と無記名評価の結果を比較した。両者の差をまとめたものを Table 5 に示す。授業評価の総合点は、記名評価が平均 92.26 点 ($SD=6.42$)、無記名評価が平均 91.87 点 ($SD=6.48$) であり、ともに最高点は 100 点、最低点は 70 点であった。総合点は無記名評価の方が若干低かったが統計的に有意な差ではなかった。授業評価を質問項目別にみると、すべての質問項目 (10 項目) において無記名評価の評定値が記名評価の評定値を下回っていたが、いずれも大きな差ではなかった (最も差が大きかった項目「授業に熱意が感じられたか」で 0.11 点)。Table 5 からわかるように、授業評価の総合点以外の項目についても、記名評価と無記名評価間で評定値に大きな差はみられなかった。

次に研究 4 と同様、無記名で行われた授業評価の総合点と各変数間の相関係数を求めた。結果を Table 6 に示す。出席状況の自己採点 ($p<.001$)、授業態度の自己採点 ($p<.001$)、試験の自己採点 ($p<.001$)、教員に対する好感度 ($p<.001$) の 4 項目と授業評価の総合点との間に有意な相関がみられ、記名評価で認められた相関関係は無記名評価においても確認された。ただし、算出された相関係数は、記名評価よりも無記名評価の方が概ね高い値となっており (研究 4 の Table 4 参照)、無記名で授業評価を求めた時の方が変数間の関係がより顕著に現れていたようである。

最終的な成績が発表されていない段階で授業評価

Table 6
授業評価の総合点と各変数間の相関係数
(無記名評価)

出席状況の自己採点	.28***
授業態度の自己採点	.30***
試験の自己採点	.22***
教員に対する好感度	.40***
成績のつけかた	-.06
試験の難易度	.03

*** $p<.001$

を記名および無記名で求め、両者の比較を行ったが、本研究の結果をみる限り、記名評価と無記名評価の間に大きな差は認められなかった。

研究 6

学生による授業評価を実施している多くの大学では、当該科目の授業評価を 1 回実施するだけであるが、同じ授業に対する評価内容は一貫して時間経過とともに変化しないものだろうか。もし変化がみられるとしたらどのように変化するのか。本研究では、こうした点を確かめるため、受講学生に同一授業の評価を時間を隔てて 2 回求め、両時期の評価内容に違いがみられるか否かについて検討した。

方法

対象者 1994 年度前期および 1995 年度前期開講の一般教育科目「心理学」を受講した女子短期大学生 505 名 (1994 年度の受講者は 253 名, 1995 年度の受講者は 252 名)。

Table 7
各評価項目に対する平均評定値

評価項目	1994 年度		1995 年度	
	1 回目	2 回目	1 回目	2 回目
①授業はわかりやすかったか	9.46	9.31	9.33	9.37
②授業はまとまっていたか	9.29	9.15	9.21	9.09
③授業の進み具合はどうだったか	9.15	8.81 ^{**}	8.95	8.74
④教員の熱意は感じられたか	8.92	8.87	8.90	8.74
⑤教員は授業の準備を十分にしていたか	9.49	9.41	9.43	9.29
⑥トピックの選び方は適切だったか	9.23	9.03	9.07	8.95
⑦板書は見やすかったか	8.53	8.17 ^{**}	8.23	8.22
⑧声の大きさはどうだったか	9.59	9.27 ^{**}	9.33	9.00
⑨話すスピードはどうだったか	9.20	8.97	9.07	8.85
⑩教材はどうだったか	9.02	8.70 ^{**}	9.07	8.86

^{*} $p < .05$, ^{**} $p < .01$

手続き 1994 年度受講者, 1995 年度受講者を対象に「心理学」の授業評価を 2 回ずつ求めた(無記名評価)。授業評価を実施した時期は, 1994 年度受講者が 1994 年 9 月と 1996 年 1 月(両時期の間隔は 16 カ月), 1995 年度受講者が 1995 年 9 月と 1996 年 4 月である(両時期の間隔は 7 カ月)。1 回目の授業評価は 1994 年度受講者, 1995 年度受講者ともに試験時に行ったため単位認定前の評価である。それに対し, 2 回目の授業評価は単位認定後の評価となる。受講者に配付した授業評価シートにはこれまでの研究と同じ 10 項目からなる授業評価項目(①「授業はわかりやすかったか」, ②「授業はまとまっていたか」, ③「授業の進み具合はどうだったか」, ④「教員の熱意は感じられたか」, ⑤「教員は授業の準備を十分にしていたか」, ⑥「トピックの選び方は適切だったか」, ⑦「板書は見やすかったか」, ⑧「声の大きさはどうだったか」, ⑨「話すスピードはどうだったか」, ⑩「教材はどうだったか」)が含まれており, それぞれの質問に対して 0 点(非常に悪い)~10 点(非常に良い)で評定してもらった。授業評価シートの提出は自由とした。

結果と考察

1994 年度受講者については 1 回目が 240 名, 2 回目は 214 名分の回答が得られた。また 1995 年度受講者については 1 回目は 252 名の回答が得られたが, 2 回目の回答は 65 名だけであった。これらのデー

タをもとに各評価項目に対する評定値を平均した。評価項目中未記入等の欠損値については除外して分析を行った。結果を Table 7 に示す。

Table 7 からわかるように, 1994 年度受講生, 1995 年度受講生ともにほとんどの評価項目において 1 回目よりも 2 回目の平均評定値が低くなっていた(2 回目に増加していたのは 1995 年度受講者の評価項目①のみ)。特に, 1994 年度の授業に対する評価では, 10 項目中 4 項目(③, ⑦, ⑧, ⑩)において 1 回目よりも 2 回目の評価が有意に低くなっていた。

前述のように, 1 回目の授業評価が単位認定前の評価であったのに対し, 2 回目の評価は単位認定後の評価である。単位認定前に授業評価を求めた場合, 全体的に評価が甘くなることも推測され, 本研究において 1 回目と 2 回目の評価内容に差がみられたのは, この点が影響していたのかもしれない。しかし, 1995 年度よりも 1994 年度の方が 1 回目と 2 回目の平均評定値間の差が大きかったことから, 時間の経過も評価内容の悪化に影響していたと考えられる。

本研究の結果からも明らかのように, 学生による授業評価の内容は時間軸上で必ずしも一定ではない。本研究で評価の対象となった授業に関しては, 時間経過とともに学生による授業評価の内容が悪化した, 逆のケース(時間経過とともに評価内容が好転すること)もあるかもしれない。いずれにせよ, 授業評価を求める時期が評価内容に影響を及ぼす可

能性は否定できない。

総合考察

本稿では、藤田らが検討した学生による授業評価に影響を及ぼすと考えられるいくつかの要因を紹介した。研究1では、試験の得点、出席状況・授業態度・試験についての自己採点、単位のとりやすさに関する認識の違いに着目し、授業評価の総合点との関係について調べた。その結果、授業評価の総合点と相関がみられたのは、試験の得点、出席状況の自己採点、授業態度の自己採点、試験の自己採点であった。研究2では、授業評価と学生の性格特性間の相関関係について検討した。学生の性格特性を測る指標としてエゴグラムを用いたが、エゴグラムの5つの尺度のうち授業評価の総合点と有意な相関がみられたのは、NP, A, FCの3尺度であった。こうした傾向が異なる科目に対する授業評価でもみられるか否かを検討したのが研究3であるが、授業評価の総合点とエゴグラムの5つの尺度の間には一貫した傾向をみいだすことはできなかった。しかし、研究2と研究3においてエゴグラムの5尺度の得点総和と授業評価の総合点の間には4科目中3科目で統計的に有意な相関が確認された。研究4では、学生が教員に対してもつ好感度と授業評価との関係について検討した。その結果、授業評価の総合点と教員に対する好感度の間には強い正の相関関係がみられた。研究5では、授業評価を記名で行う場合と匿名（無記名）で行う場合で授業評価の結果を比較したが、両者の間に統計的に有意な差はなかった。研究6では、学生による授業評価の内容が時間の経過とともにどのように変化するのかについて検討した。授業評価の対象となった科目に関しては、ほとんどの評価項目において1回目よりも2回目の評価の平均評定値が低くなっており、学生による授業評価の内容は時間軸上で必ずしも一定ではないことがわかった。

ほとんどの大学では、授業の質的改善を学生による授業評価の目的としている。しかし、学生による授業評価に対する批判の多くはこの点に関係したものである。すなわち、「学生による授業評価は本当に授業改善のために役立っているのか」、「授業の改善を学生による授業評価の目的にしている大学は、その目的がどの程度達成されているのか検証しているのか」といった批判である。こうした批判は、学生による授業評価について述べられた文献の中で頻繁にみられるが、学生による授業評価をもとに授業の改善を試みるかどうかは教員個人に委ねられてお

り、実際に授業の質的改善がみられたかどうかを判断することは極めて難しい。“授業の質”は多角的に判断されるべきものであり、授業評価の結果だけで“授業の質”を判断することは危険である。

アメリカの大学では、学生による授業評価を、教員の昇進や昇給、あるいは終身在職権（テニユア）獲得の際の1つの判定資料として用いる場合もある（喜多村, 1990）。日本においても、学生による授業評価の結果をもとに「優秀授業科目」を選出して表彰したり、授業評価の結果を“ベストティーチャー賞”の選出基準の1つとして利用したりする大学もある（江本, 2010；小田, 2010；田口, 2007）。しかし、“授業の質”に関する明確な定義や基準がないまま、学生による授業評価の結果だけを“授業の質”を測る指標としてとらえ、教員の教育能力を査定する道具として用いるのは、学生による授業評価に対する過信であり、誤用であると思われる。

中央教育審議会は2008年12月24日の答申「学士課程教育の構築に向けて」の第3章「学士課程教育の充実を支える学内の教職員の職能開発」の中で学生による授業評価について触れている。そこでは「学生による授業評価の結果は、業績評価の指標としての信頼性には課題もあるが、教員の自己評価や職能開発の活動に生かすことは重要であると考えられる」という見解が示されており、学生による授業評価の結果を短絡的に教員の人事考査に利用することに疑問符をつける一方で、学生による授業評価は教員が自らのティーチング・スキルを磨く上で活用するには有用であるとしている。学生による授業評価について長年研究を続けている米谷（2010）は学生による授業評価について、「基本的には授業評価は健康診断のようなものであり、問題がなければそれでよいのであり、それで序列をつけたりするのはいかなものかと思う」としているが、現状ではこうしたとらえ方が妥当ではないと思われる。

本研究の結果から示唆されるように、学生による授業評価の内容は様々な要因によって変動し、授業の内容や形式以外の要因によっても評価の内容が左右される可能性がある。また、学生のレベルや授業の中で学生がなにを求めているかによっても授業評価の内容は異なってくるのが考えられる。そのため、学生による授業評価を有効に活用するためには、実際にその大学で行われた授業評価の特徴を詳細に分析することが必要である。本稿で紹介した一連の研究は1990年代に行われたものであるが、学生による授業評価に影響を及ぼす諸要因がその後大きく変わったとも思われぬ。本学における学生による

授業評価を実りあるものとするためには、こうしたケーススタディーを積み重ね、その結果を教員間で共有することが必要だと思われる。

引用文献

- 阿部桂子 (1987). 学生による講義評価と予想成績および実際の成績との関係. *Niigata educational psychologist*, 4, 1-2.
- Algozzine, B., Beattie, J., Bray, M., Flowers, C., Gretes, J., Howley, L., Mohanty, C., & Spooner, F. (2004). Student evaluation of college teaching: A practice in search of principles. *College Teaching*, 52, 134-141.
- Centra, J.A. (1993). *Reflective faculty evaluation: Enhancing teaching and determining faculty effectiveness*. San Francisco, CA: Jossey-Ball.
- 中央教育審議会 (2005). 新時代の大学院教育 (答申)
- 中央教育審議会 (2008). 学士課程教育の構築に向けて (答申)
- デュセイ, J. M. (1980). エゴグラム: ひと目でわかる性格の自己診断 池見酉次郎 (監修) 新里里春訳 創元社
- 江本理恵 (2010). 根拠に基づいた教育改善活動を目指して—授業アンケートから授業の公開・共有に向けて— 東北大学高等教育開発推進センター編 学生による授業評価の現在 東北大学出版会 pp.209-216.
- 藤田勉・川島眞 (1993). 大学生による講義評価に影響を及ぼす諸要因について 日本応用心理学会第60回大会発表論文集, 172-173.
- 藤田勉・久東光代・川島眞 (1994). 大学生による講義評価に影響を及ぼす諸要因についてⅡ 日本応用心理学会第61回大会発表論文集, 94.
- 藤田勉・西村誠・清水道夫・小林洋文・久東光代・川島眞 (1995). 大学生による講義評価に影響を及ぼす諸要因についてⅣ 日本教育心理学会第37回総会発表論文集, 107.
- 藤田勉・久東光代・川島眞 (1995). 大学生による講義評価に影響を及ぼす諸要因についてⅤ 日本応用心理学会第62回大会発表論文集, 30.
- 藤田勉・久東光代・川島眞 (1996). 大学生による講義評価に影響を及ぼす諸要因についてⅦ 日本応用心理学会第63回大会発表論文集, 29.
- 藤田勉・久東光代・川島眞 (1997). 大学生による授業評価に関する研究 日本人間関係学会第5回大会発表要旨集, 50-51.
- 井上理 (2010). 「学生による授業評価」調査の活用を考える 東北大学高等教育開発推進センター編 学生による授業評価の現在 東北大学出版会 pp.65-106.
- 川島眞・久東光代・藤田勉 (1994). 大学生による講義評価に影響を及ぼす諸要因についてⅢ 日本応用心理学会第61回大会発表論文集, 95.
- 川島眞・久東光代・藤田勉 (1995). 大学生による講義評価に影響を及ぼす諸要因についてⅥ 日本応用心理学会第62回大会発表論文集, 31.
- 川島眞・久東光代・藤田勉 (1996). 大学生による講義評価に影響を及ぼす諸要因についてⅧ 日本応用心理学会第63回大会発表論文集, 30.
- 喜多村和之 (1990). 大学淘汰の時代—消費社会の高等教育 中央公論社
- 喜多村和之 (1992). 大学評価とはなにか—ア krediteーションの理論と実際 東信堂
- 米谷淳 (2010). 大学教育改革と授業評価 東北大学高等教育開発推進センター編 学生による授業評価の現在 東北大学出版会 pp.5-36.
- Marsh, H.W. (1987). Students' evaluation of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, 11, 253-388.
- Milgram, S. (1965). Some conditions of obedience and disobedience to authority. *Human Relations*, 18, 57-76.
- 文部科学省高等教育局大学振興課大学改革推進室 (2013). 大学における教育内容等の改革状況等について (平成23年度)
- Naftulin, D.H., Ware, J.E., & Donnelly, F.A. (1973). The doctor fox lecture: A paradigm of educational seduction. *Journal of Medical Education*, 48, 630-635.
- 新里里春 (1986). 基本的構えとエゴグラム 新里里春・水野正憲・桂戴作・杉田峰康 交流分析とエゴグラム チーム医療 pp.1-22.
- 小田隆治 (2010). 公開・共有化を目指した学生と教員による授業改善アンケート 東北大学高等教育開発推進センター編 学生による授業評価の現在 東北大学出版会 pp.217-229.
- 大塚雄作 (2010). 授業評価の読み方・使い方—学問学習共同体における実践的妥当性のすすめ— 東北大学高等教育開発推進センター編 学生による授業評価の現在 東北大学出版会 pp.37-64.
- 大山泰宏 (2007). 授業評価の発想と歴史 山地弘起 (編著) 授業評価活用ハンドブック 玉川大学出版部 pp.11-30.
- 田口真奈 (2007). 授業評価の諸機能 山地弘起 (編著) 授業評価活用ハンドブック 玉川大学出版部 pp.31-51.
- 竹綱誠一郎 (1992). 大学生の講義評価と出席度およびテスト得点との関係. 日本心理学会第56回大会発表論文集, 394.
- 谷口守 (2005). 授業評価に基づくティーチング技術アップ法 技報堂出版
- Wachtel, H.K. (1998). Student evaluation of college teaching effectiveness: A brief review. *Assessment & Evaluation in Higher Education*, 23, 191-211.
- 山本晴義 (1990). 人間関係ゲーム 廣済堂出版
- 安岡高志 (2005). 学習の質・量を充実させるために 大阪大学大学院原子力工学専攻編 学びに成功する「よい授業」とは何か 大阪大学出版会 pp.110-127.

Zimbardo, P.G. (1970). The human choice: Individuation, reason, and order versus deindividuation, impulse, and chaos. In W.J. Arnold & D. Levine(eds.) *Nebraska Symposium on Motivation*, University of Nebraska Press.

Zimbardo, P.G., Haney, C., Banks, W.C., & Jaffe, D. (1977).

The Psychology of imprisonment: Privation, power and pathology. In J.C. Brigham & L.S. Wrightman(eds.) *Contemporary Issues in Social Psychology*. 3rd ed. Cole Publishing Company.

(平成 26 年 10 月 1 日受付、平成 26 年 11 月 28 日受理)