

学生および教員自身の授業評価はどの程度一致するか？

林 創

(岡山大学大学院教育学研究科)

To What Extent Are Course Evaluations by Students and that by the Teacher Similar for a Class?

Hajimu Hayashi

(Graduate School of Education, Okayama University)

Summary

The purpose of this study is to examine to what extent two types of evaluations are similar: course evaluations by students and that by the teacher for the same class. Seventy-nine students who took a particular class and the teacher of the class were required to evaluate each lecture throughout one semester. The results showed that there was a difference between the variances of students' evaluations and the variances of the teacher's evaluations over the semester. However, the tendencies of both evaluations were similar for many individual lectures. Furthermore, a significant correlation was found between the two types of evaluations. These findings indicate that course evaluations by the students might be generally analogous to the evaluations by the class's teacher.

キーワード：授業評価, 形成的評価, 学生, 教員

Keywords: course evaluation, formative evaluation, students, teacher

1. 問題と目的

近年、日本の大学は大きな変革を遂げている。1991年の大学設置基準の「大綱化」とともに、大学は「自己点検・自己評価」が義務づけられ、大学自身に「評価」が課せられるようになった(大塚、2005)。この流れは21世紀に移るころから加速化し、2002年の中央教育審議会の答申を受けて学校教育法が改正され、「大学の質の保証」に関して、文部科学大臣に認証された第三者評価機関による「認証評価」を、すべての大学が7年に1度ずつ受けることが義務づけられたのである。

このような大学評価の広がりのもとには、大学教育そのものの改善が強く求められるようになったという背景がある。いわゆるFD(Faculty Development)と呼ばれるものがその中核を成す。FDとは多様な解釈がされる用語であるが、おおまかに「大学教員の個人的・集団的な教育力の養成活動」と理解されている(田中、2003)。1998年の大学審答申によって、大学の第三者評価やFDの必要性などが強調され、同時期に大学評価・学位授与機構が設置された。さらには、2006年の大学院設置基準の改正、および2007年の大学設置基準の改正により、大学院および学部での「FDの法制的義務化」という新たな事態に直面(田中、2008)し、大学は対応を求められるようになった。大学教育と大学評価の両者の流れが相俟って、日本の大学においてFD活動は定着することになったのである。

そのようなFD活動の一環として、最も広く普及しているものの一つが、学生による「授業評価」である。現在では、ほとんどの大学で何らかの形で実施されており、授業の改善の判断材料として使われている。たとえば、関西地区の大学を対象とした授業評価の現状を調べた調査(林・大塚、2008)からは、「全学的に実施している」と回答した大学が98%にもものぼっていた。さらに、全科目で実施(66%)し、回答者が特定されないような配慮がなされ(90%)、項目数は11~20項目(63%)程度で構成され、5段階評定(73%)とし、自由記述欄を設け(98%)、回答方法とし

てマークシートを利用 (77%) することで集計の負担を減らし、試験前の最後の授業 (40%) で1回だけ実施している (92%) というものが平均的な授業評価の実施方法であることが明らかになっている¹⁾。

このような調査からわかるように、授業評価の多くは、半期の授業で終わりの方に1回だけ実施されている。しかしながら、教育評価の実施にはいくつかの方法がある。たとえば、実施時期に着目すると、ある教育プログラムが始まる前に学習者がどのような知識や技能を身につけているかを評価する「診断的評価」、教育プログラムが実施されている途中で学習者がどの程度目標を達成しているかを把握する「形成的評価」、教育プログラムが終了した後に成果を評価する「総括的評価」の3種類に区分される。この3種類の区分は、もともとスクリヴァン (Scriven, M.) が提唱した形成的評価と総括的評価の着想に学んで、ブルーム (Bloom, B. S.) がまとめたものとして知られている (田中、2005)。この枠組みで考えると、半期の授業で終わりの方に1回だけ実施されるということは、平均的な大学の授業評価が、総括的評価に相当する形式で実施されていることを意味する。

総括的評価が大切であることに異論はないであろうが、授業改善という視点を考えた場合、形成的評価も重要であると考えられる。なぜなら、形成的評価をすることで、指導を学習者の状態にあわせて適切に変化させることができるからである。もう少しかたい言葉でいえば、「教授活動を (できるだけ客観的に) モニターし、授業ごとに、即応的に修正や改善といったコントロールを行うことができる」という、より良いダイナミズムが期待できるからである (北神・林、2005)。

実際に、形成的評価を重視し、毎回の授業評価を行うことで、翌週以降の授業の進捗や目標を即応的に改善する試みも報告され、授業の向上に有益であることが示唆されている (e.g., 藤田、2005; 大塚、2005)。また、毎回の授業評価をすると、「最終回の評定だけ高い」結果が生じたことも報告されており、授業の最後だけ評価をするのでは、正確な評価ができていない可能性も示唆されている (藤田、2000)。これも、形成的評価の重要性を示す一つの知見といえるだろう。このように、形成的評価を踏まえた授業は効果的であり、毎回の授業で、感想用紙を配布し、形成的評価をする意味があると考えられる。

ところが、形成的評価を含めた授業を実際に進めていると、教員自身の授業に対する感覚と学生の授業評価にギャップがあると感じられる場面もある。たとえば、教員自身が準備を十分に行って授業に臨み、「今日は良い授業ができたぞ!」と手ごたえを感じた回であっても、学生の反応があまり良くない (例: 理解度が低い、関心がない) ように感じることもある。また、逆に教員自身が「今日の授業はうまくいかなかった」と落ち込んでいた回であっても、学生のその回の授業評価が意外にも高かったと感じることがある。このような経験は、形成的評価を行っていない場合でも、大学教員であれば誰しも直観的に何度か感じたことがあるのではないだろうか。一般に、授業が半期15回という長期にわたって続くものであることを客観的に考えれば、このようなギャップが何回かは生じるのは当然といえるかもしれないが、教員自身が主観的に感じるギャップは、そのような客観的な考えを超えるレベルで頻繁に生じているようにも思われる。

そこで本研究では、半期の授業の各回で、学生の授業評価に加えて教員自身の授業評価を実施し、両者の評定結果にどの程度の一致やギャップがあるのかを検討することを目的とした。その際、単に学生と教員の双方が授業に対する評価を独立に行うだけでなく、それぞれ相手がどのような評価をするかを想像する状況も試みた。すなわち、学生は教員がどう評価するかを推測してもらい、逆に教員は学生が (全体として平均的に) どう評価するかを推測し、評定項目に対する回答を求めた (図1)。これらの間の評定結果を比較することにより、同じ授業に対する教員の感じ方と学生の感じ方のズレを多面的に認識することができ、より良い授業を構成していく糸口が探れるはずである。

2. 方法

調査対象者 岡山大学教育学部で「学校教育心理学」(平成20年度前期・2単位)を受講した1年生79人、および、担当教員(筆者)が調査の対象者であった。

授業日程 調査の対象となった「学校教育心理学」の授業日程を表1に示す。出張により休講が1度あり、授業回数は計14回であった。このうち、学生の約半数が5月23日に、残りの約半数が6月13日に、附属学校での教育実習に参加したため、本授業には出席できなかった。そのため、5月23日と6月13日の授業内容は同一であった。授業は、教科書『絶対役立つ教育心理学』(藤田、2007)に沿って進められた。

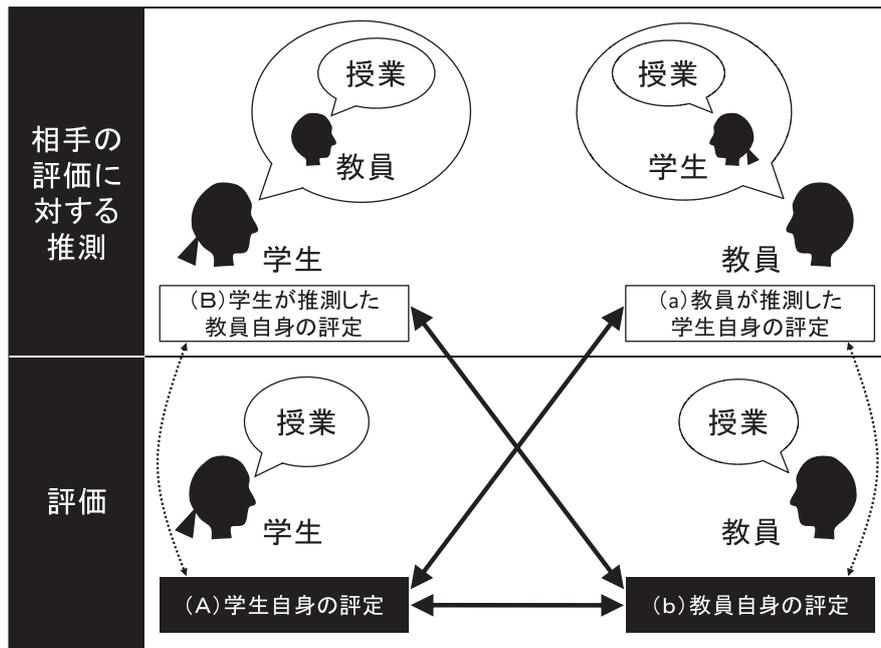


図1 本研究での4種類の評価とその関係

表1 授業日程

第1回	(4/11)	講義 (学問としての教育心理学の紹介)
第2回	(4/18)	講義 (学習のメカニズム)
第3回	(4/25)	講義 (動機づけの基礎)
第4回	(5/9)	講義 (動機づけの応用、記憶の基礎)
第5回	(5/16)	講義 (記憶の応用)、DVD 視聴
第6回	(5/23)	講義 (学習方略1)
第7回	(5/30)	講義 (メタ認知と学習観)
第8回	(6/6)	講義 (発達の理論)
第9回	(6/13)	講義 (学習方略2)
第10回	(6/20)	講義 (教育の応用)
第11回	(6/27)	講義 (児童期と青年期の教育)
第12回	(7/4)	講義 (障害の理解)
第13回	(7/11)	期末試験と答え合わせ
	(7/18)	出張のため休講
第14回	(7/25)	授業の総括、テストの総評

表2 評価項目

受講生回答版：(A) 今日の授業について、あなた自身の評価をして下さい (教員回答版：(a) 今日の授業について、受講生はどう評価したと思うか?)
1. 今日の授業での教員の話し方 (聞き取りやすさ)
2. 今日の授業での教員の熱心さ
3. 今日の授業のわかりやすさ
4. 今日の授業内容に興味を持てたか、刺激を受けたか
5. 今日の授業に対するあなたの理解度の自己評価
受講生回答版：(B) 教員は、今日の授業をどう評価すると思いますか? (教員回答版：(b) 今日の授業についての自己評価)
1. 今日の授業での教員の話し方 (聞き取りやすさ)
2. 今日の授業での教員の熱心さ
3. 今日の授業でのわかりやすさ
4. 受講生は今日の授業内容に興味を持てたか、刺激を受けたか
5. 受講生の今日の授業に対する理解度の自己評価

調査用紙と手続き 藤田 (2005) を参考に、毎回の授業終了時に、その回の授業について「教員の授業の仕方」と「学生の自己評価」を問う項目で構成された感想用紙を作成した。その際、表2のように、「(A) 学生自身の評価項目」(5項目)と、「(B) 教員はどう評価するかを推測させる項目」(5項目)の2つを用意した。(A)と(B)の各項目は、同一の概念を問うものとなり比較できるように、可能な限り表記を揃えた。

学生には、出席の確認を兼ねて、(A)と(B)が印刷された感想用紙を配布し、各項目に対して「非常に良かった = 5」から「非常に悪かった = 1」までの5段階で評定してもらった²⁾。

他方、教員(筆者)は、毎回の授業終了後に、あらかじめ表計算ソフト(MicrosoftのExcel 2007)上に用意しておいた(A)と(B)に対応する「(a) 学生はどう評価するかを推測する項目」と「(b) 教員自身の評価項目」に対して、(b)については、学生79人が平均的にどのような評定になるかを推測しながら学生と同様に5段階で評定した。その際、学生の評定の影響を受けないように、回収された学生の感想用紙を見る前に評定を済ませた。

3. 結果と考察

3.1 授業の各回の評定の推移

半期で14回実施された授業の各回で、学生が評定した項目である「(A) 学生自身の評定」と「(B) 学生が推測した教員自身の評定」、および教員自身が評定した項目である「(a) 教員が推測した学生自身の評定」と「(b) 教員自身の評定」の4つのそれぞれについて、5項目の評定の個人ごとの平均値を算出した。さらに、学生が評定した「(A) 学生自身の評定」と「(B) 学生が推測した教員自身の評定」については、79人の平均を算出した。その結果を、図2に示す。

14回にわたっての内的整合性を検討するために α 係数を算出したところ、学生が評定した項目である「(A) 学生自身の評定」と「(B) 学生が推測した教員自身の評定」、および教員自身が評定した項目である「(a) 教員が推測した学生自身の評定」と「(b) 教員自身の評定」の4つのそれぞれについて、.87、.93、.91、.93であり、いずれも十分な値であった。なお、「(A) 学生自身の評定」と「(B) 学生が推測した教員自身の評定」の内的整合性を検討する際は、それぞれ各回の授業での5項目について79人の平均値をデータとみなして、14回にわたっての α 係数を算出した。

図2に示された14回を通しての授業評価の変化から、学生の評定の変動が小さいことが明らかになった。学生の評定は、「(A) 学生自身の評定」と「(B) 学生が推測した教員自身の評定」のいずれにおいても、14回を通して、おおむね4.0~4.5の間を推移した。これに対して、教員の評定は、「(a) 教員が推測した学生自身の評定」と「(b) 教員自身の評定」のいずれにおいても、2.5~4.5のあたりで、授業の回ごとに大きな変動が見られた。とくに、「(b) 教員自身の評定」が3.0以下であった回が、14回のうちの半数近くの6回で見られたことから、教員自身の評価は低

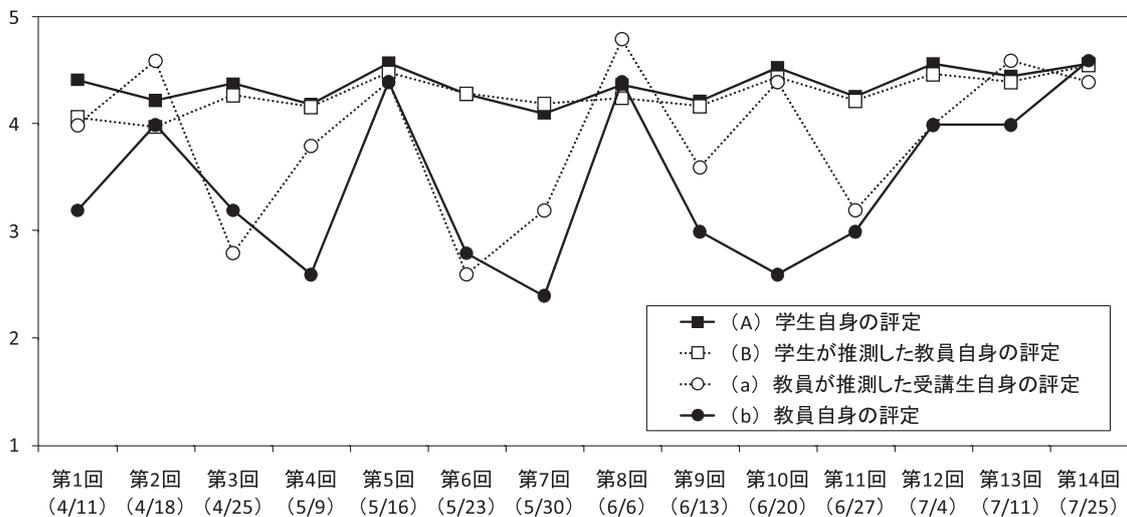


図2 各回の評定の平均値の推移

めの値であった。

以上を踏まえると、5段階評定の「絶対値」で比較した場合、変動の大きさの違いにより、学生と教員の評価の関係がわかりにくくなる。それゆえ、ここでは、「(A) 学生自身の評定」と「(b) 教員自身の評定」について、第1回(初回)と第14回(最終回)を除く12回について、各回の前後との変動の方向性で比較を行うことにした。

その結果、「学生と教員の変動の方向性はかなり一致する」ことが明らかになった。具体的には、第4回から第8回と第12回から第13回の合計7度の授業において、「(A) 学生自身の評定」の前後の変動の方向性と「(b) 教員自身の評定」のそれが一致していた³⁾。たとえば、第4回の前後の変動に着目すると、第3回から第4回にかけて、(A)は下降していたが、(b)も下降しており、第4回から第5回にかけて、(A)は上昇していたが、(b)も上昇していた。他方、「(A) 学生自身の評定」の前後の変動の方向性と「(b) 教員自身の評定」のそれが完全に逆だったのは、第2回と第10回の2度のみであった。この場合、第1回から第2回にかけて、(A)は下降していたが、(b)は上昇しており、第2回から第3回にかけて、(A)は上昇していたが、(b)は下降していた。また、第9回から第10回にかけて、(A)は上昇していたが、(b)は下降しており、第10回から第11回にかけて、(A)は下降していたが、(b)は上昇していた。

ただし、本研究では、教員は1人のデータに対して、学生は79人のデータであり、学生の評定の変動が小さくなるのは統計学的にも当然と考えられる結果である。上記のように、教員の(平均)評定値と学生全員の(平均)評定値の平均値を比較するのは、次元がそろっておらず、必ずしもフェアではない。そこで、各回の授業での5項目の評定の個人ごとの平均値を用いて、さらにすべての回(14回)にわたっての平均値と標準偏差を、教員を含めて個人ごとに算出した⁴⁾。その結果、教員のすべての回にわたっての平均値は3.44で、標準偏差は0.73であった。他方、学生のすべての回にわたっての平均値は79人のうち最大値が4.96、最小値が3.40、中央値が4.40で、標準偏差は79人のうち最大値が0.71、最小値が0.11、中央値が0.37であった。学生の平均値の最小値と教員の平均値が同程度であったことから、教員の評価は低めの値であったことが追認された。また、学生の標準偏差の最大値よりも教員の標準偏差の方が大きかったことから、比較の次元を揃えた分析からも、教員と学生の変動に差があり、総じて学生の方が変動が小さいことが追認されたと考えられる。

さらに、学生の標準偏差の最大値と最小値の差が大きく、評定の個人差が大きいため、その差を踏まえた上で、教員の評定の推移とどの程度違いがあるのかを検討することにした。具体的には、すべての回にわたっての評定の標準偏差が大きい学生X(標準偏差は0.70)、中央値の学生Y(標準偏差は0.37)、小さい学生Z(標準偏差は0.16)を取り出して、教員の評定と比較した。なお、学生Xと学生Zは、教育実習による欠席1回を除く13回の授業すべてに出席した学生の中で、標準偏差の値が最も大きい学生と最も小さい学生であり、それを基準に選んだ。また、学生Yも教育実習による欠席1回を除く13回の授業すべてに出席した学生であった。3人の学生と教員について、各回の評定の平均値の推移を図3に示す(3人の学生のいずれもが第6回の授業日に教育実習に参加し、本授業には出席できなかったため、図3では第6回を省いている)。

分析の結果、3人のいずれについても「学生と教員の変動の方向性はかなり一致する」傾向があった。学生Xでは第4回から第8回と第11回から第13回の合計7度の授業において、学生Yでは第4回から第7回と、第12回から第13回の合計5度の授業において、学生Zでは第2回と第4回から第8回と第13回の合計6度の授業において、「(A) 学生自身の評定」の前後の変動の方向性と「(b) 教員自身の評定」のそれが一致していた⁵⁾。たとえば、第4回の前後の変動に着目すると、第3回から第4回にかけて、3人の学生とも(A)は下降していたが、(b)も下降しており、第4回から第5回にかけて、3人の学生とも(A)は上昇していたが、(b)も上昇していた。他方、「(A) 学生自身の評定」の前後の変動の方向性と「(b) 教員自身の評定」のそれが完全に逆だったのは、学生Xの第2回と第10回のみであった。この場合、たとえば第1回から第2回にかけて、(A)は下降していたが、(b)は上昇しており、第2回から第3回にかけて、(A)は上昇していたが、(b)は下降していた。

以上のように、前後の回との「変動の方向性」という点に注目すると、学生と教員の授業評価はかなり一致することが明らかになった。

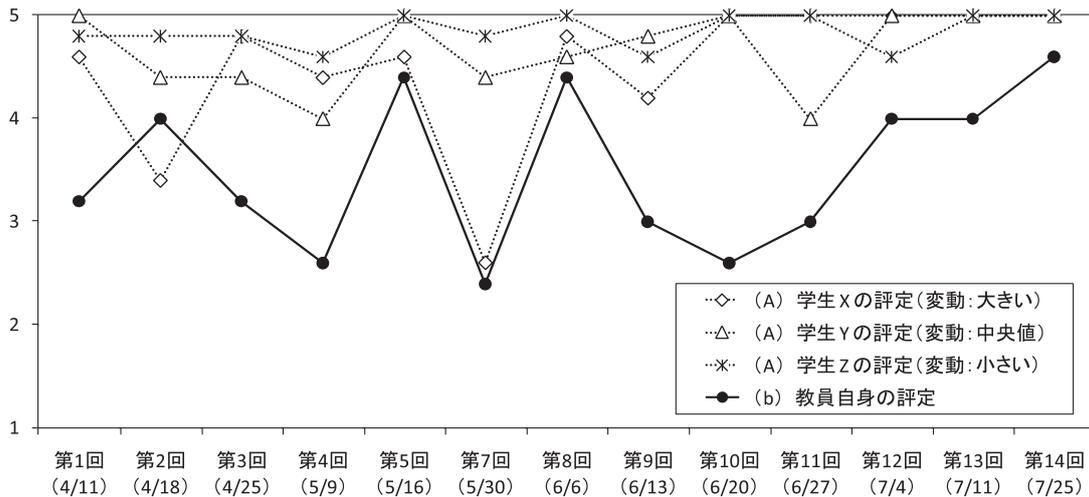


図3 各回の評定の平均値の推移 (第6回は学生のデータがないため除外)

3.2 評定間の相関

次に、各評定間の相関を算出することで、さらに深い分析を行うことにした(図4)。

まず、14回の授業を通しての「(A) 学生自身の評定」と「(b) 教員自身の評定」の間に有意な正の相関があった($r=.61, p<.05$)。このことから、教員がうまく進めることができたと感じた授業は、学生も肯定的にとらえている傾向が高く、教員があまりうまくできなかったと感じた授業は、学生も相対的に低くとらえている傾向が見られた。

また、「(A) 学生自身の評定」と「(a) 教員が推測した学生自身の評定」の間には、10%水準での有意な相関($r=.47, p<.10$)が見られたが、「(B) 学生が推測した教員自身の評定」と「(b) 教員自身の評定」の相関は有意でなかった($r=.40, ns$)。両者の相関係数の差の検定を行ったところ、有意な差はなく、教員と学生のいずれも、授業に対する相手側の感じ方を把握する程度に差はなかった。

さらに、「(A) 学生自身の評定」と「(B) 学生が推測した教員自身の評定」の間 ($r=.78, p<.01$)、および「(b) 教員自身の評定」と「(a) 教員が推測した学生自身の評定」の間 ($r=.67, p<.01$) に有意な相関があった。このことは、それぞれが「他者の視点」に立つことを求められても、「自分自身と似た評定になる」傾向が高いことを示すと考えられる。この結果を踏まえると、独立している「学生自身の評定」と「教員自身の評定」の間の相関係数 .61 は、比較的高い値であると解釈することができるであろう。

4. 全体的考察

本研究の目的は、半期の授業の各回で、学生の授業評価に加えて教員自身の授業評価を実施し、両者の評定結果にどの程度の一致やギャップがあるのかを検討することであった。

まず、授業の各回の評定において、前後の回との「変動の方向性」という点に注目すると、学生と教員の授業評価はかなり一致することが明らかになった(図2)。さらに、各評定間の相関を算出(図4)したところ、14回の授業を通しての「(A) 学生自身の評定」と「(b) 教員自身の評定」の間に有意な正の相関があったことから、教員が高く評価した授業は、学生も相対的に高めに評価している傾向があり、教員が低く評価した授業は、学生も相対的に低めに評価している傾向が見られた。

以上より、教員の「今日の授業はうまくできた」や「今日の授業はダメだった」という直観的感覚は、少なくとも本研究では、学生全体が感じる授業への肯定的もしくは否定的な評価とおおむね一致する傾向があったと言えるだろう。この結果は、教員が主観的に感じやすいギャップ、すなわち、教員自身の授業に対する感覚と学生の授業評価の間の違いを比較的頻繁に感じやすいことと矛盾する結果である。

この矛盾は、変動の大きさの違いによって生み出される部分があるかもしれない。結果でも述べたように、教員と学生の変動に差があり、総じて学生の方が変動が小さかった。それゆえ、教員の感覚の変化の大きさ(たとえば、「前回よりかなり良い授業ができたぞ!」という感覚)に、多くの学生の感覚の変化の大きさがついてこず、あまり変化

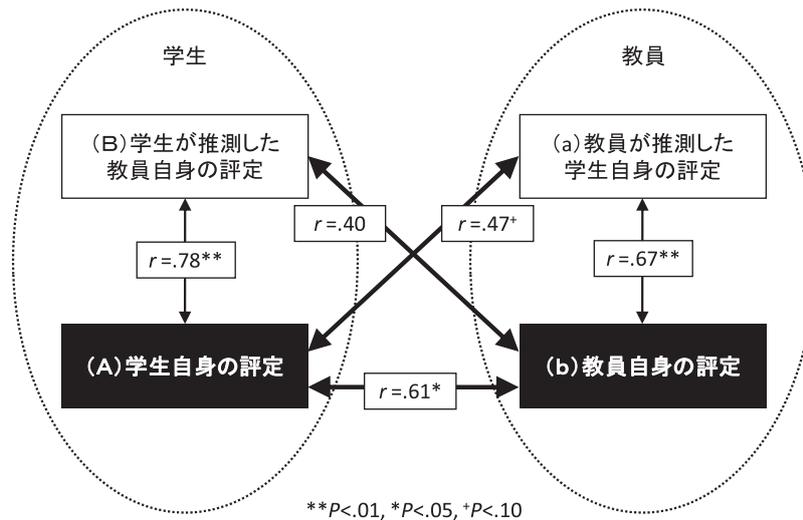


図4 評定間の相関

がなかったように感じられてしまうのかもしれない。

さらに、上記の矛盾を説明する1つの可能性として、人間の「認知」の特性から迫ることができるかもしれない。近年、「認知心理学」と「行動経済学」の研究が大きく進展し、人間は決して合理的な思考をせず、さまざまな歪みを生む考え方をすることが次々と明らかになっている。こうした研究では、ある対象を評価する際に、自分の感情や願望に左右されたり、対象の目立ちやすい部分に焦点が合い、他の特徴についての評価が歪められる現象を「認知バイアス」と呼ぶ。

認知バイアスは様々な形で生じるものである。たとえば、「風呂に入っていると、電話がかかってくる気がする」といったことや、「自分が行きたい方向へ行くエレベーターやバスを待っていると、逆方向へ行くエレベーターやバスばかり先に来る気がする」といったことは、多くの人々が日常的に感じていることである (Gilovich, 1991)。もし入浴中に電話がかかってきたとすると、そのことは注目され、強く印象に残って記憶にされやすい。なぜなら、濡れたままあわてて電話口に出て、寒い思いをしながら話したり、せっかく風呂場から出てきたのに、電話が切れてしまって腹を立てたりといったことが起こるからである。逆に、入浴中に電話がかかってこなかった場合には、それが出来事として記憶されることはない。別に何も起こっていないので、特別な感情を引き起こすこともないからである (Gilovich, 1991)。このように、ネガティブな出来事は、実際にはそれほど生じていなくても、心理的な面で蓄積されやすいという傾向がある。

本研究の結果も、この心理的傾向によって、少なくとも一部を説明できる可能性がある。多くの教員にとって、「事前に周知の準備をして、うまく進めた授業において、学生の評価が低かった」ということは、強いネガティブ経験となる。そのため、実際にはそれほど回数でなくとも、ネガティブな経験が強く印象に残って記憶されてしまい、その結果、教員自身の授業に対する感覚と学生の授業評価にギャップがあると思ってしまう可能性が考えられる。したがって、本研究のような客観的な方法で授業評価を分析すると、心理的に思っていたほどには学生と教員の間で評価に差が無く、実際以上にネガティブに感じていたことに初めて気づくのもかもしれない。

ただし、このような認知バイアスは、授業改善を考えた時にプラスの方向に働くと思われる。なぜなら、教員が「実際以上に授業がうまくいかない」と主観的に感じやすい傾向は、授業に対する不満の感覚を解消するために、授業をより良いものにしようという動機を頻繁に生じさせやすくさせ、その結果、授業改善につながる面があると考えられるからである。それゆえ、このような認知バイアスは、現実を歪めてとらえる「悪いもの」ではなく、むしろ「良いもの」と考えることができるであろう。

本研究からは以上のような知見が導かれた。ただし、このような解釈を進める上で、いくつかの注意をしておく必要がある。

第1に、教員の授業への肯定的もしくは否定的な評価は、学生全体が感じる授業への肯定的もしくは否定的な評価

とおおむね一致する傾向がみられたものの、相関係数が十分に高いわけではなかった。それゆえ、実際に授業を進める際は、「今日の授業がうまくできた」と思った回が多いとしても、学生の評価が完全に連動するわけではないことを意識し、授業に対する改善の視点を忘れないようにすべきであろう。そのためには、形成的評価を重視した毎回の授業評価の実施が有益と思われる。

第2に、本研究では1つの授業での結果をまとめたにすぎない点に注意すべきである。複数の教員での授業、あるいは1人の教員の場合であれば複数の授業で、同じような傾向が見られるのかを確認する必要があるだろう。授業評価は、講義型と演習型という授業形態によっても違いが生まれる(澤田、2006)ため、授業形態の影響を受けるかどうかも検討すべきである。また、その際は、心理学関連の授業だけでなく、違う学問分野、とくに理系の授業などでも同様の結果になるかを調べる必要がある。このような一般化への過程を踏まえて、本研究の知見が確かなものかを確認することが可能になる。その過程では、個々の授業は文脈や背景などが違うため、自らの実践の中で評定平均値を意味づけていく「実践的妥当化」(大塚、2005)もますます重要になってくると思われる。

最後に、今後の課題を考えてみたい。第1に、本研究の結果を説明する1つの枠組みとして「認知バイアス」の影響を検討したが、認知バイアスは様々なものが確認されている。そのような多様性を考えれば、授業の改善につながる認知バイアスだけでなく、逆方向に作用する認知バイアスも考えられるかもしれない。その際は注意を促すこともできるであろう。認知心理学的な視点から授業評価を考えることは、これまであまりなかったと思われ、大学教育の研究を新たな角度から深めていく1つの方向性となるかもしれない。

第2に、本研究における「(B) 学生が推測した教員自身の評定」と「(a) 教員が推測した学生自身の評定」は、学生と教員がそれぞれ「相手がどのように感じるか」を推測して回答するものであった。これは、心理学において「心の理論」という概念と関連することである(e.g. 林、2010)。心の理論とは、簡単に言えば、相手の行動や言葉を見たり聞いたりして、相手の考えや気持ちを察することである(坂井、2008)。心の理論は、通常の大人であれば、ふつうに機能していることであるが、それがどの程度深いレベルで働いているかは個人差があると考えられる。同じ状況に置かれても、深く相手のことを考える人もいれば、相手のことをあまり考えず、自分が思うままの人もいるであろう。それゆえ、今回の評定に際しても、学生79人の個人差がいろいろあったと考えられる。より精密な研究をするためには、たとえば、授業以外の場合に、心の理論がどの程度機能するのかを、個人ごとに調べておくべきなのかもしれない。このような問題には、社会心理学の研究の知見も援用できることであろう。

第3に、学生は各自が、教員1人に対して、どのように感じているかを考えることになるが、教員は1人で、学生が「全体(本研究では79人)として、平均的にどのように感じているか」を考えることになり、学生と教員の間で、推測のしかたに非対称性が生じている。今後は、このような点も考慮した上で、分析を深めていかなければならない。

以上のような注意点や課題はあるものの、心理学などの学問的知見を生かしながら研究を進めていくことで、FD活動の一環として普及している「授業評価」を、さらに有益なものとしていくことができるであろう。

謝 辞

本論文にまとめる以前に、第15回大学教育研究フォーラムや、岡山大学教育学部「授業研究会」などで、本研究の発表をお聞きいただきました皆様から貴重なご助言とご示唆を賜りました。記して感謝申し上げます。

註

- 1) この調査(林・大塚、2008)は、「関西地区FD連絡協議会」(調査時は準備会)が企画した「授業評価ワークショップ」の開催(2008年1月12日)に合わせて実施されたものである。関西地区の大学213校にアンケートが送付され、授業評価ワークショップへの参加希望を表明した大学54校が分析対象となった。
- 2) 感想用紙は、上半分が(A)と(B)の評定項目の欄、下半分が授業に対する感想や質問を自由に記述する欄で構成されていた。
- 3) 厳密には、第12回から第13回にかけて、「(b) 教員自身の評定」は、4.0から4.0へのままで変動がなかった。

ので、完全には方向性が一致しているとは言えないが、繁雑になるため、ここでは方向性が一致していることに含めた。

- 4) 「方法」で記したように、学生の約半数が5月23日に、残りの約半数が6月13日に、附属学校での教育実習に参加したため、本授業には出席できなかった。さらに、それ以外の回で休んでいる学生も存在した。そのため、学生についての「すべての回」とは、最大で13回であった。すべての回にわたっての平均値と標準偏差も、各自の出席した回のデータだけを用いて算出した。
- 5) 脚注3)と同様に、繁雑さを避けるため、数値に変動がなかった部分も合わせて、ここでは方向性が一致していることに含めた。

引用文献

- 藤田哲也 2000 「学生の受講態度の自己評価と授業評価との関係について」『光華女子大学研究紀要』38号、249-268頁。
- 藤田哲也 2005 「動機づけ理論をふまえた授業運営」 溝上慎一・藤田哲也(編著)『心理学者、大学教育への挑戦』ナカニシヤ出版、79-114頁。
- 藤田哲也(編著) 2007 『絶対役立つ教育心理学—実践の理論、理論を實踐—』ミネルヴァ書房。
- Gilovich, T. 1991 *How we know what isn't so: The fallibility of human reason in everyday life*. New York: Macmillan Publishing Company. (守一雄・守秀子訳 1993『人間この信じやすきもの—迷信・誤信はどうして生まれるか—』新曜社。)
- 林 創 2010 「心の理論—他者の『心』を理解する心」 森 敏昭・淵上克義・青木多寿子(編)『よくわかる学校教育心理学』ミネルヴァ書房、122-123頁。
- 林 創・大塚雄作 2008 「関西地区FD連絡協議会『授業評価ワークショップ』事前アンケートとその結果」京都大学高等教育研究開発推進センター(編)『関西地区FD連絡協議会設立に向けて』、133-149頁。
- 北神慎司・林 創 2005 「高水準リテラシー教育を實踐する上でのメタ認知の重要性」 溝上慎一・藤田哲也(編著)『心理学者、大学教育への挑戦』ナカニシヤ出版、217-222頁。
- 大塚雄作 2005 「学習コミュニティ形成に向けての授業評価の課題」 溝上慎一・藤田哲也(編著)『心理学者、大学教育への挑戦』ナカニシヤ出版、1-37頁。
- 坂井克之 2008 『心の脳科学—「わたし」は脳から生まれる—』中公新書。
- 澤田忠幸 2006 授業評価の年次変化と授業タイプによる違いの影響『大学教育学会誌』28号、102-109頁。
- 田中耕治 2005 「教育評価の機能：診断的評価、形成的評価、総括的評価」 田中耕治(編著)『よくわかる教育評価』ミネルヴァ書房、8-9頁。
- 田中每実 2003 「ファカルティ・ディベロップメント論—大学教育主体の相互形成」 京都大学高等教育研究開発推進センター(編)『大学教育学』培風館、87-106頁。
- 田中每実 2008 「はじめに」 京都大学高等教育研究開発推進センター(編)『関西地区FD連絡協議会設立に向けて』、1-4頁。