

## 授業評価調査における中間評価の有効性

細川 和仁

### Is the Midterm Survey of Class Evaluation Effective in the Improvement of the Quality of Classes?

Kazuhiro HOSOKAWA

#### 1. 問題

##### 1.1. 授業評価の現状

本研究は、大学の授業評価における「中間評価」が、授業改善にどのような影響を与えているかについて検討し、その有効性を明らかにすることが目的である。

大学における教育内容等の改善については、カリキュラム改革や授業方法の開発をはじめとして、成績評価方法や履修科目数の上限設定、組織的な研修活動（FD）など多岐にわたって行われている。学生による授業評価も、それらの取り組みの一つである。文部科学省の調べによれば、平成16年度の時点で学生による授業評価を実施している大学は、国立大学では100%、公立大学・私立大学ではそれぞれ97%となっており、ほとんどの大学で何らかの形で導入されているという状況にある。

日本の大学における授業評価は、1980年代から、先導的な実践例が蓄積されてきた（沖，2004）。1990年代に入ると、体育や語学関連の科目における取り組みが多く報告されるようになり、授業を評価するという取り組み自体が広がりを見せ始めたが、全学的な実施はここ数年の動向だといえるだろう。それまでは、多くの大学関係者にとって、大学の授業は何らかの評価の対象として見られることはなかった。「あの授業はおもしろい／つまらない」と学生が感じるがあっても、それはあくまでインフォーマルな評価であり、それ以上でもそれ以下でもなかった。授業を担当する教員

は、カリキュラムを逸脱しない範囲で授業の計画・実施及び学生の成績評価に対して責任を持ち、その営みに対して、他者が意見を述べるようなことは行われてこなかった。

このことは、大学教員が授業を疎かにしていたことを意味するものではない。ここで指摘すべきことは、授業に関する仕事が個々の教員の個業によって行われてきた点である。いわゆる「オムニバス形式」の授業のように、1つの科目を複数の教員で担当することはあっても、複数の教員が共同で授業設計をする風土はなく、授業方法について学ぶ場もない。また、教員同士で授業を見合うような習慣もない。よって、よりよい授業づくりに取り組む教員がいたとしても、それを教員間で共有するような仕組みはない。教員は、自分自身が受けてきた授業の経験に基づいて、試行錯誤しながら授業を実施していく以外に方法がなかったといえる。

しかし、大学教育のあり方が問い直され、教育改善に取り組むことが大学に求められる中で、組織的な取り組みとしての授業評価も急速に普及した。学生による授業評価に対しては、依然として「やる気のない学生や基礎学力のない学生に授業を評価する資格はない」といった反対意見もある。その一方で、6割程度の教員は学生の授業評価に対して一定程度の信頼を置いているという調査結果も報告されている（安岡，2004）。学生による授業評価が、授業改善に向けた取り組みへの契機の一つになっていることは間違いなさそうである。

## 1.2. 授業評価と授業改善

授業評価が授業改善の契機となっている点は、義務教育における授業改善とは異なり、大学に特徴的な状況だといえる。小学校や中学校における授業改善は、教員に対する研修という形で進められてきた経緯がある。教員の授業実践力を高めるため、教育委員会や教育センター等が研修を企画し、各学校においては「校内研修」という形で、授業の実践的な研究が積み重ねられている。児童・生徒による授業評価が導入されている小・中・高等学校は、まだごく一部である。

大学においては学生による授業評価が定着したが、やがて課題として浮上してくるのはマンネリ化、形骸化であろう。アンケートを実施すればそれでよしという時代は終わり、授業評価の結果をいかに授業改善に結び付けているかが問われるようになってきている。現在わが国の大学で行われている授業評価は、教員やカリキュラム評価ではなく、授業の改善や指導の効果を知らることが目的とされており、授業評価は自らの授業実践を振り返るための一つの道具とみなすことができる。つまり、自らの授業実践にどのような課題があるのかを知る手がかりである。授業評価の結果から課題を見いだすことができれば、それを「次の授業」に生かすことができる。逆に言えば、それが可能になるような授業評価でなければならない。

授業評価に関するこれまでの研究としては、授業評価尺度の因子構造や、各項目の関連性についての分析が多く見られる。例えば、田中・藤田(2003)は、達成目標と授業自体の評価、受講態度の自己評価、学業遂行との関係を検討し、相関を明らかにしている。また、星野・牟田(2003, 2005)は、授業評価項目の因子分析により、評価・教授努力・コミュニケーションの3因子を抽出し、それらの関連性を分析している。

一方、授業評価結果と授業改善に関わる教育活動の関連性にまで言及した研究としては、澤田(2006)や鈴木(2006)の研究が挙げられる。澤田は、4年間にわたって授業評価とそれに基づく改善活動を行い、授業形態によってどのような効果の差異があるかを検討している。また、鈴木は、自身が担当する授業科目について、授業評価の結果を受けて次年度に何を改善したのか、それが授業評価結果にどう反映されたかについて、5年に

わたって継続的に調査し、それに加えて、履修した学生に追跡調査を実施し、多角的な分析を行っている。

## 1.3. 授業評価における中間評価

授業評価を授業改善に生かすことを考えたとき、評価を授業のどの時点で実施するかは重要なポイントとなる。沖(2004)は、授業評価を実施時期の点から3つに類型化している。すなわち、授業の事前(最初)、中間、事後(最終)の3つに分類できるというものである。現在行われている授業評価の多くは事後(最終)評価である。事後評価は、授業の最終に行われるため、回答した結果が直接的に回答者に還元されることはない。教員にとっても、調査から得たことは、次年度に開講されるその授業に生かされることになる。

事後評価の限界を克服し、授業評価の結果を授業改善に生かすには、中間評価が有効な方法となりうる。沖(2004)は中間評価にも2種類あるとしている。一つは、毎回の授業終了時に行う、記述形式の簡単なアンケートである。「大福帳」「レビューシート」「ミニッツ・ペーパー」と呼ばれるものがこれにあたり、授業の中間段階での評価に位置づけることが可能である。もう一つが、授業の導入部分が終了した時期に実施するアンケートである。なお大山(2003)はこれを「early evaluation」と呼び、授業の4, 5回目で実施し、残り10回の授業の改善に用いる、と指摘している。本論文でいう中間評価は、この授業の中間段階でのアンケートを指し、その有効性について論じる。

授業の中間段階での評価の実施状況は、東北大学が行った調査から傾向を把握することができる(東北大学高等教育開発推進センター, 2006)。東北大学が実施した調査は、全国大学教育研究センター等協議会の加盟大学を対象に実施したもので、国立大学の現状をある程度写し出したものといえる。その結果によれば、平成17年度の時点で「中間評価」を実施している大学はまだ少なく、茨城大学(工学部)、愛媛大学(共通教育, 理学部)、鹿児島大学(共通教育)のみである。よって、この中間評価の取り組みは、現時点では先進的なものといえる。

東北大学の調査の対象には含まれていないが、徳島大学の全学共通教育における実施例が報告さ

れている（松谷・桑折・佐野，2006）。徳島大学の全学共通教育では，中間アンケートを実施した上で，学期末の授業評価アンケートにおいて，「授業は改善されたか」及び「中間アンケートは授業改善に効果があると思うか」という項目を設け，中間アンケートの評価を直接学生に尋ねている。その結果は，肯定的な評価の割合が否定的な評価を上回ってはいるが，「どちらともいえない」という回答も依然として多く，学生の多くが判断がつかねる状態にある，としている。

秋田大学の教養基礎教育においては，平成16年度2期からこの中間評価が導入されている。評価をその後の授業の改善に生かすという機能に焦点を当て「形成的評価」と呼ばれている。

この中間評価が，授業改善にどのような影響を及ぼしているのかを明らかにすることは，授業評価と授業改善の結びつきを確認する点で意義があると考えられる。また，それらの結びつきが弱いという結果が得られれば，授業改善に資するような評価のあり方について検討することもできる。以上の観点から，本論文では，秋田大学教養基礎教育科目の授業評価調査における，中間評価（形成的評価）と最終評価（総括的評価）の関係を明らかにすることを目的とする。

（註）この後の記述では，中間評価を「形成的評

価」，最終評価を「総括的評価」と記載する。

## 2．研究の手続き

### 2.1．対象科目

対象とするのは，平成17年度に開講された本学の教養基礎教育科目のうち，「形成的評価」を実施した科目の授業評価調査結果である。形成的評価実施の対象となっているのは，15週にわたって開講する科目である。形成的評価は授業の7週目または8週目に実施し，事務担当者が集計したのち，その結果はすみやかに担当教員に報告されることになっている。担当教員以外には結果は公開されていない。また，「総括的評価」は，授業の最終回（15週目）に実施されている。

平成17年度の対象科目は1期241科目，2期143科目，あわせて384科目であった。そのうち形成的評価と総括的評価の両方の調査が行なわれた科目は，1期151科目（実施率62.7%），2期82科目（同57.3%），計233科目（同60.7%）であった。調査が実施された科目の各科目における回答数は表1の通りである。調査に回答した学生の数とその科目の登録者数，単位認定者数等は必ずしも一致しないと考えられるが，本論文では暫定的に，調査回答者数をクラスサイズと置き換えて考えることとする。

表1：本論文で対象とする授業科目のクラスサイズ（単位：%）

回答数	1～20	21～40	41～60	61～80	81～120	121～	計 (N)
1期	12.6	23.8	32.5	17.2	10.6	3.3	100.0 (151)
2期	12.2	20.7	31.7	14.6	13.4	7.3	100.0 (82)
全体	12.4	22.7	32.2	16.3	11.6	4.7	100.0 (233)

### 2.2．質問項目

本稿で対象とする質問項目は，形成的評価と総括的評価の質問項目のうち，両方の評価で設定されている項目である（表2参照）。具体的には，目的や目標が明確に説明されているか，授業がよく準備されているか，授業に対する熱意が感じられるか，授業内容が興味深いものか，説明が明確で十分にわかりやすいか，学生の理解度に配慮した形で進められていたか，という教員の授業のやり方に対する項目に加えて，授業の内容が十分に身に付いたか，という内容の習得

に対する学生の自己評価の計7項目である。これらの質問項目に対して，「そう思う」「どちらかといえばそう思う」「どちらともいえない」「どちらかといえばそう思わない」「そう思わない」の5つからあてはまるものを選ばせる。以下の結果と考察においては，「そう思う」= 5，以下4，3，2，1と数値に置き換え，平均値を求めている。この平均値は，例えば受講者全員が「どちらかといえばそう思う」と答えれば，その科目の平均値は4.00となる。

表2：本論文で対象とする質問項目

	形成的評価	総括的評価
①目的や目標の明確な説明	1.授業の目的や目標が明確に説明されていましたか。	Q3-1.授業の目的や目標が明確に説明されていましたか。
②授業の準備	2.授業はよく準備されていましたか。	Q3-2.授業はよく準備されていましたか。
③教員の熱意	3.授業に対する教員の熱意が感じられましたか。	Q3-3.授業に対する教員の熱意が感じられましたか。
④興味深い内容	4.授業の内容や進め方は興味深いものでしたか。	Q3-4.授業の内容は興味深いものでしたか。
⑤明確でわかりやすい説明	5.説明は明確で十分にわかりやすいものでしたか。	Q3-6.説明は明確で十分にわかりやすいものでしたか。
⑥学生の理解度への配慮	6.授業は学生の理解度に配慮した形で進められていたと思いますか。	Q3-7.授業中の指示や助言等を含め、学生の理解度に配慮した進め方でしたか。
⑦内容の習得に対する自己評価	7.授業の内容が十分に身に付きましたか。	Q3-8.授業の内容が十分に身に付きましたか。

### 3. 結果と考察

#### 3.1. 形成的評価・総括的評価の比較

はじめに、形成的評価、総括的評価の結果を確認しておこう。

全233科目の各科目の学生による評価の平均値を割り出し、さらにその平均値を求めたのが表3である。形成的評価では、授業の準備や教師の熱意に関しては、平均値が4.00を超えており、比較的高い値だといえよう。ついで、目的や目標の明確な説明が3.90となっている。以下、興味深い内容、明確でわかりやすい説明、学生

の理解度への配慮が続き、内容の習得に対する自己評価が最も低く、平均3.38となっている。一方、総括的評価は、形成的評価と同様、項目と項目が平均4.00を超えており、その他の項目も同じような数値の傾向を示している。

形成的評価と総括的評価の平均値を比較すると、すべての項目で総括的評価が上回っている。2回の評価の間に統計的に有意な差があるかどうかについてt検定を行ったところ、設問～の項目については有意な差は見られなかったが、～の項目については有意差が見られた。

表3：形成的評価と総括的評価の各項目の平均値

	形成的評価	平均値の差	総括的評価
①目的や目標の明確な説明	3.90		3.93
②授業の準備	4.01		4.03
③教員の熱意	4.06		4.07
④興味深い内容	3.65	*	3.77
⑤明確でわかりやすい説明	3.60	*	3.68
⑥学生の理解度への配慮	3.55	*	3.67
⑦内容の習得に対する自己評価	3.38	*	3.53

自由度はいずれも232、\*は $p < 0.05$ で有意差があることを示す。

全体平均の差とあわせて、各科目の比較の結果も見ておく。表4は、各科目の形成的評価と総括的評価の値を比較し、どちらが高い値になっているかについて科目の割合を求めたものである。形成的評価<総括的評価、すなわち最終評価で評価が伸びた科目は、～の項目では全体の約5割強にとどまっているが、項目～は全体の3分の2程度、項目～では7割を超えている。

この結果から、次のようなことがいえる。まず、すべての項目で総括的評価の値が形成的評価の値を上回ったことから、これらの項目の観点については、学生から最終的により良い評価を得られるようになっていることがわかる。しかし、目的や目標の明確な説明、授業の準備、教員の熱意の3項目に関しては統計的な有意差は見られず、総括的評価でより良い評価を得た科目も半数

表4：形成的評価と総括的評価の比較（科目の割合）（単位：％）

	形成的評価＞ 総括的評価	形成的評価＝ 総括的評価	形成的評価＜ 総括的評価	計 (N)
①目的や目標の明確な説明	43.3	2.6	54.1	100.0 (233)
②授業の準備	46.4	1.3	52.4	100.0 (233)
③教員の熱意	46.8	1.7	51.5	100.0 (233)
④興味深い内容	31.3	2.1	66.5	100.0 (233)
⑤明確でわかりやすい説明	35.6	1.3	63.1	100.0 (233)
⑥学生の理解度への配慮	32.6	1.3	66.1	100.0 (233)
⑦内容の習得に対する自己評価	26.6	1.7	71.7	100.0 (233)

強にとどまった。その理由として考えられることは、この3項目は、形成的評価において既に高い値を示していることである。受講者全員が最高評価「そう思う」をつけた場合、平均は5.00となるが、20人以下の少人数授業であっても、平均が5.00となることはほとんどないといってよい。4.00を超える平均値はかなり高い値だと言える。ある程度まで数値が高くなると伸び幅が小さくなる傾向が、ここに表れていることが考えられる。

### 3.2 . 形成的評価の高低による総括的評価の伸びの差異

これを検証するため、形成的評価と総括的評価

の値を、もう少し細かく見ておきたい。形成的評価が高い場合、伸びがあまり見られなくなるという仮定を検証するため、形成的評価の各項目の平均値によって科目を分類し、総括的評価の伸びを比較する。表5は、形成的評価の各項目の平均値を0.5ごとに区切り、それぞれのカテゴリーにあてはまる科目が、総括的評価においてどれだけ伸びているか、その平均値をまとめたものである。表の左上のセル「0.04 (22)」を例にすれば、形成的評価の質問の平均値が4.50以上だった科目は、全233科目中22科目あり、その22科目の総括的評価の平均値は形成的評価のそれと比較すると0.04のマイナスになっている、という意味である。

表5：形成的評価の値別にみた総括的評価の伸び（カッコ内は該当する科目数）

形成的評価 の平均値	①目的や目 標の明確な 説明	②授業の準 備	③教員の熱 意	④興味深い 内容	⑤明確でわ かりやすい説 明	⑥学生の理 解度への配 慮	⑦内容の習 得に対する 自己評価
4.5～5.0	▲0.04 (22)	**▲0.08 (37)	*▲0.06 (52)	*0.12 (15)	▲0.02 (15)	▲0.02 (15)	▲0.18 (1)
4.0～4.5	0.01 (81)	▲0.02 (89)	0.01 (80)	0.02 (52)	0.01 (51)	0.05 (37)	0.05 (25)
3.5～4.0	0.04 (92)	0.04 (75)	**0.06 (78)	**0.12 (73)	0.02 (69)	**0.08 (76)	**0.14 (69)
3.0～3.5	0.09 (32)	**0.14 (26)	0.05 (17)	**0.15 (66)	**0.15 (64)	**0.17 (71)	**0.13 (100)
～3.0	0.10 (6)	0.14 (6)	0.06 (6)	**0.32 (27)	**0.23 (34)	**0.29 (34)	**0.31 (38)
全体	0.03 (233)	0.01 (233)	0.02 (233)	**0.13 (233)	**0.08 (233)	**0.12 (233)	**0.15 (233)

該当する科目の形成的評価と総括的評価の平均値の差の検定（t検定）を行った結果、\*\*は  $p < 0.01$ 、\*は  $p < 0.05$  であることを示す。

この結果によると、形成的評価で4.5を超える科目は、以外の6項目でマイナスとなっており、総括的評価で伸びが見られない。また、4.0～4.5の科目においても、ほとんど伸びは見られない。伸びが見られたのは、形成的評価の値が4.0以下だった科目、すなわち「3.5～4.0」「3.0～3.5」「3.0以下」の3つのカテゴリーにあてはまる科目であった。特に3.0以下のカテゴリーに該当する科目の平均値を比較すると、興味深い内容については2.66～2.98、明確でわかりやすい説明は2.57～2.80、学生の理解度への配慮は2.60～2.88、内

容の習得に対する自己評価は2.68～2.99となっており、大きな伸びを見せている。

項目別に見ると、3.5～4.0、4.0～4.5、4.5～5.0は4.0以下の3つのカテゴリーで統計的に有意な伸びが見られ、項目も3.5以下の2つのカテゴリーで有意差が見られた。形成的評価と総括的評価の全体的な比較において、3.0～3.5の項目に有意な差を生じさせたのは、形成的評価ではそれほど評価が高くなかった科目の伸びにあるといえ、これらの科目の授業改善に反映されていることがうかがえる。

形成的評価の値が高い場合伸びがあまり見られ

ないという点については、表5のまとめから読み取ることができるが、形成的評価が4.0以下だった場合であっても、項目によっては伸びが確認できない場合もある。

### 3.3. クラスサイズによる伸びの差異

次に、クラスサイズによる差異について検討する。クラスサイズは授業評価結果に大きな影響を及ぼすと考えられているが、授業改善に対する効果という観点から見ると、どのような差異があるのだろうか。ここではクラスサイズによって科目を6つのカテゴリーに分け、それぞれの平均値を算出した。表6-1は形成的評価の平均値、表

6-2は総括的評価の平均値、そしてそれらの差をとって総括的評価における伸びを示したものが表7である。

まず、表6-1、6-2を見ると、クラスサイズによる値の差が確認できる。しかし、少人数であるほど評価が高まり、大人数になるほど評価が低くなるという単純な傾向は見られない。20人以下の小規模のカテゴリーは、どの項目においても評価がひじょうに高い。21~40人のカテゴリーが、その次に評価が高いカテゴリーとなっている。しかし、大規模のクラスになるほど評価が低くなるという傾向にはなっていない。20人以下を除く5つのカテゴリーは、カテゴリー間での差が小さい。

表6-1：クラスサイズ別の形成的評価の平均値

回答者数 (N)	①目的や目標の明確な説明	②授業の準備	③教員の熱意	④興味深い内容	⑤明確でわかりやすい説明	⑥学生の理解度への配慮	⑦内容の習得に対する自己評価
1~20 (29)	4.266	4.46	4.49	4.24	4.20	4.13	3.85
21~40 (53)	3.96	4.03	4.12	3.66	3.63	3.58	3.42
41~60 (75)	3.83	3.94	4.008	3.55	3.48	3.46	3.30
61~80 (38)	3.79	3.85	3.91	3.44	3.40	3.36	3.24
81~120(27)	3.77	3.93	3.89	3.57	3.49	3.42	3.28
121~ (11)	3.86	4.012	3.95	3.58	3.55	3.45	3.24
全体	3.90	4.01	4.06	3.65	3.60	3.55	3.38

表6-2：クラスサイズ別の総括的評価の平均値

回答者数 (N)	①目的や目標の明確な説明	②授業の準備	③教員の熱意	④興味深い内容	⑤明確でわかりやすい説明	⑥学生の理解度への配慮	⑦内容の習得に対する自己評価
1~20 (29)	4.272	4.40	4.55	4.31	4.19	4.16	3.92
21~40 (53)	4.01	4.04	4.17	3.82	3.76	3.74	3.59
41~60 (75)	3.86	3.97	4.011	3.69	3.57	3.62	3.50
61~80 (38)	3.78	3.88	3.86	3.55	3.49	3.46	3.38
81~120(27)	3.81	3.95	3.90	3.71	3.56	3.53	3.40
121~ (11)	3.90	4.013	3.99	3.68	3.60	3.55	3.34
全体	3.93	4.03	4.07	3.77	3.68	3.67	3.53

次に、それらの差をまとめた表7について見ていきたい。平均値の差の検定を行ったところ、21~40人及び41~60人のカテゴリーでは、～の項目で有意な差が見られた。また、61~80人、81~120人のカテゴリーでも、項目、において有意差が見られた。その一方、20人以下の小規模のカテゴリー、及び121人以上の大規模のカテゴリーは、有意な差が見られなかった。また、～、～の質問項目については、どのクラスサ

イズにおいても優位な差が見られないことが明らかになった。

これらの結果については、次のような要因が考えられる。まず、20人以下のカテゴリーは、形成的評価、総括的評価の両方において、その他のどのクラスサイズよりも、すべての項目において高い評価を得ており、これ以上大きく伸びることは考えにくいということがあげられる。少人数のクラスは、中間評価の段階ですでに高い評価を得て

表7：クラスサイズ別の総括的評価の伸び

回答者数(N)	①目的や目標の明確な説明	②授業の準備	③教員の熱意	④興味深い内容	⑤明確でわかりやすい説明	⑥学生の理解度への配慮	⑦内容の習得に対する自己評価
1～20 (29)	0.01	▲0.06	0.06	0.07	▲0.01	0.03	0.07
21～40 (53)	0.05	0.01	0.05	**0.16	*0.13	**0.16	**0.17
41～60 (75)	0.03	0.03	0.00	**0.14	**0.09	**0.16	**0.20
61～80 (38)	▲0.01	0.03	▲0.05	*0.11	0.09	*0.10	**0.14
81～120(27)	0.04	0.02	0.01	**0.14	0.07	**0.11	*0.12
121～(11)	0.04	0.00	0.04	0.10	0.05	0.10	0.10
全体	0.03 (233)	0.01 (233)	0.02 (233)	0.13 (233)	0.08 (233)	0.12 (233)	0.15 (233)

t検定により，\*\*は  $p<0.01$ ，\*は  $p<0.05$  であることを示す。

おり，その評価を最終段階まで持続させているともいえる。また，121人以上の大規模のクラスに有意な伸びが見られなかったのは，このカテゴリーに含まれる科目数が少なかったこと（11科目）が原因の1つだと考えられる。平均値では若干ではあるがすべての項目で伸びが見られている。

### 3.4．考察

これらの結果から，数値だけで見れば，どの項目においても，総括的評価は形成的評価よりも向上している。しかし，詳しく見ていくとその向上の程度は一樣ではない。伸びが著しいのは，質問項目という観点からは，興味深い内容，明確でわかりやすい説明，学生の理解度への配慮，内容の習得に対する自己評価，の4項目であった。

項目～の評価の向上は，形成的評価において4.0未満だった科目，クラスサイズの観点からは，小規模（20名以下）でも大規模（121名以上）でもない科目の評価が伸びていることによる。逆に言えば，これらに該当する科目の改善に，形成的評価は役立っているといえる。

その他の3項目，すなわち目的や目標の明確な説明，授業の準備，教員の熱意は統計的に有意な伸びが見られなかった。その理由として，形成的評価において既に，3.90を超える高い評価を得ていることが原因と考えられた。しかし，表5に見られるように，目的や目標の明確な説明については，形成的評価の高低にかかわらず，統計的に有意な伸びは見られない。このことから，  
・ ・ の3項目は，形成的評価の結果を受けて総括的評価に向けた授業改善に生かすことが難しい観点であり，授業の序盤に評価がある程度固

定してしまう可能性が高い項目である。言い換えれば，授業の担当教員がこれらの観点を特に意識すべきなのは，授業の序盤であるといえる。授業の目的や目標を明確に説明し，周知な準備をした上で授業にのぞみ，それらを通じて，授業に対する教員の熱意を学生に伝えることが，授業開始後の数回において特に重要になると言えよう。教員の熱意などは，授業の途中から変化する（ように学生が見る）ことは考えにくく，むしろ興味深い授業内容にすること，明確でわかりやすい説明をすること，学生への理解度に配慮した進め方をすることなどは，その改善状況を学生は，授業の全般に渡って敏感に感じ取っているといえる。

また，クラスサイズの点から見ると，小規模（20名以下）の授業や大規模（121名以上）の授業で大きな伸びが見られなかった点は，形成的評価が効果がなかったことを示すと断じることはできないだろう。20名以下のクラスサイズの授業は，どの観点においても，他のどのクラスサイズの場合よりも有意に高い評価を得ていた。よって，形成的評価の時点で既に高い評価を得ていることで，総括的評価の伸びが見られなかったと解釈することが可能である。

## 4．授業改善に結びつけるために

### 4.1．形成的評価の授業改善への影響

中間評価が授業改善に与える影響という観点からは，秋田大学教養基礎教育で実施しているものに関して言えば，一定の良い影響を与えているとみることができる。特に，興味深い内容，明確でわかりやすい説明，学生の理解度への配慮に大きな改善が見られ，そのことが学生の内容の習得に対する自己評価の向上につながっていることが示

唆された。また、クラスサイズという点から見れば、20人以下の小規模のクラスや121人以上の大規模のクラスでは大きな改善は見られなかったものの、それ以外の場合では授業改善につながっていることが明らかになった。そして、形成的評価では4.0以下だった授業科目に関して、総括的评价で大きな伸びが見られたことは、全体的な底上げに機能していることもわかった。

#### 4.2. 今後の課題

しかし、このような方法の授業評価が、どのような場合でも授業改善に有効に機能するわけではない。今後の課題として、次の点を検討していく必要がある。

##### 形成的評価の項目

形成的評価に盛り込まれている評価項目は、本稿で対象とした7項目のほかに、学習態度についての自己評価、教員が自由に設定できる項目、及び自由意見欄(「授業や教員に対する意見や感想、授業の改善してほしい点や優れた点など、自由に書いてください」)がある。評価項目は、この他にも加える必要はないだろうか。

星野・牟田(2005)は、先行研究の整理を通じて、授業の満足度に影響を与える要因を4種類に分類している。具体的には、教員側の要因(教授努力や学生、授業との関わり方など)、学生側の要因(授業に対する努力の度合い、やる気の種類、体調など)、授業のインプット(授業に対する動機、学生の学力・既習知識、施設設備)、

授業のアウトプット(理解度、達成度、成績)の4つである。この例を参考にしつつ、新たな質問項目の開発を考えることも、意義あることと思われる。

また、実状として自由意見欄があまり活用されていないことも課題である。評価観点に対する数値評価では拾いきれない学生の意見を集める場として自由意見欄が設けられているが、現時点では記入数が少ない。記入のための時間を確保したうえで、自由意見欄への記入を促す必要があるだろう。平均値として数値で表された結果よりも、自由意見欄に記された学生の直接的な意見が参考になることも多い。授業における教員と学生は「サービスを提供する者/受ける者」の関係ではあるが、授業を共に作っているという点では協働関係

にある。学生も授業改善の担い手であることを意識させ、建設的な意見を記入できるよう指導していく必要がある。

##### 学生への説明

授業評価の課題として、回答した結果が学生にフィードバックされていないという課題も残されている。冒頭に紹介した徳島大学の取り組みは、この点について参考にすべきものである。徳島大学の取り組みは、中間評価が授業改善に生かされたかどうかを、最終評価において率直に学生に尋ねるものであるが、このような質問項目を評価に加えることで、学生は中間評価が授業改善の道具として用いられていることを意識できる。調査を実施する上で基本的なこととして、回答者に対して調査のねらいを的確に伝えることが必要である。

##### 総括的评价の「形成的機能」

本学の教養基礎教育では、中間評価を形成的評価、最終評価を総括的评价と呼んでいるが、総括的评价にも「形成的機能」を持たせる必要がある。つまり、総括的评价の結果を「次の授業」の改善につなげるためのシステムが必要である。しかし、大学授業は「次の授業」までの時間が長い。例えば、4月～7月に15週にわたって開講された授業は、次に開講されるのは次年度の4月からである。7月に実施した授業評価の結果を、9ヵ月後の授業に生かすためには、何らかの仕組みを開発することが必要である。

以上、今後の課題として3点を取り上げたが、授業評価全体にわたる課題として、アンケートという手法についても再検討する必要があるだろう。マークシート等の方法を用いてはいるものの、授業評価には膨大なデータ処理がついてまわる。授業評価の効果と対照させながら、適正な実施規模を見極める必要があるだろう。

#### 引用文献

- 星野敦子・牟田博光(2003)大学生の授業評価にみる受講者の満足度に影響を及ぼす諸要因、『日本教育工学会論文誌』27(Suppl.) pp.213-216
- 星野敦子・牟田博光(2005)大学の授業における諸要因の相互作用と授業満足度の因果関係、『日本教育工学会論文誌』29(4), pp.463-473
- 片岡徳雄・喜多村和之編(1989)『大学授業の研究』, 玉川大学出版部
- 松谷 満・桑折範彦・佐野勝徳(2006)授業方法に関



- する中間アンケートの効果分析 授業評価の  
新たな試みと課題 , 徳島大学 『大学教育研  
究ジャーナル』第3号, pp.30-43
- 南 学 (2003) 単位の認定・不認定の予告が授業評  
価に与える影響, 『大学教育学会誌』25(2),  
pp.68-74
- 森 和彦・佐々木典彰・古川芳子 (2000) 教授法の改  
善を目的とした学生による授業評価システム  
に関する一考察, 『秋田大学教養基礎教育研究  
年報』2, pp.15-25
- 沖 清豪 (2004) 授業評価と指導法の改善, 児玉善  
仁・別府昭郎・川島啓二編著 『大学の指導法  
学生の自己発見のために』東信堂,  
pp.135-152
- 大山泰宏 (2003) 大学教育評価論, 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, pp.39-62
- 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(1),  
pp.102-109
- 鈴木 誠 (2006) 学ぶ意欲を引き出す授業とは何か 2  
授業評価のフィードバックによる授業改善  
, 北海道大学 『高等教育ジャーナル 高等  
教育と生涯学習』14, pp.99-116
- 田部井 潤・飯田考充 (2003) 第三者による授業評価  
分析の試み 学生参加型授業を事例として ,  
『大学教育学会誌』25(2), pp.119-125
- 田中あゆみ・藤田哲也 (2003) 大学生の達成目標と授  
業評価, 学業遂行の関連, 『日本教育工学会論  
文誌』27(4), pp.397-403
- 東北大学高等教育開発推進センター (2006) 『「学生に  
よる授業評価」実施状況の調査と新たな「授  
業評価改善システム」構築に向けて・報告と  
提言』
- 安岡高志 (2004) 授業改革の実際 学生による授業評  
価の視点 , 絹川正吉・舘 昭編 『学士課程  
教育の改革』, 東信堂