

学生による授業評価の課題と展望

澤田 忠幸*

Reconsideration of the Course Evaluation by Students for its Better Utilization

Tadayuki SAWADA

Key Words : 多面的授業評価, 測定論, 顧客満足度分析, 教授・学習, 学生の心理特性

序 文

Students' evaluations of teaching effectiveness, like the teaching that they intended to represent, should be viewed as a multidimensional construct.

e.g., A teacher may be quite well organized but lack enthusiasm.

Marsh, H.W.¹⁾

今日, 学生による授業評価アンケート (course evaluation : 以下, 授業評価と略す) を実施している大学は, 国公立を問わずほぼ100%に達している^{2) 註1)}。また, 以前では授業評価を実施している場合でも, 項目別平均点などの集計結果を担当教員に返却するだけであることが多かったが, 今日では授業評価を有効に活用するためのさまざまな組織的取り組みも行われ始めている。例えば, 授業改善計画書の作成と一体化した運営³⁾, ファカルティ・ディベロッパー (Faculty Developer ; FDer) による教員個々に対するコンサルテーションの実施⁴⁾⁵⁾, ベスト・ティーチャー賞に代表される教員表彰 (teaching awards) などはその一例といえる。また, 国際基督教大学などのように, 学生の評価能力自体を養成する取り組みを行っている大学もある⁶⁾。

その一方で, 早くから授業評価を実施してきた大学では, 授業評価導入による初期の効果は薄れ, 実施の形骸化や効果への疑問の声も囁かれ始めている⁷⁾。その背景には, かつてのような授業評価尺度の信頼性や妥当性に対する疑念, あるいは学生の評定能力に対する感情的不信とは異なり, 授業評価の結果をどのように解釈し, どのような理論的枠組みに基づいて授業評価を活用すればよいのかに関する戸惑いや混乱があるように感じられる⁸⁾⁹⁾。

そこで, 本稿では, 評価測定論および教育学の観点か

ら, 授業評価研究が明らかにしてきた知見を振り返り, 授業評価の効用と限界を踏まえつつ, 有効に活用していく方向性について考察することを目的とする。

授業評価をめぐる今日の問題 : 「改善」と「評価」

今日, 一般的に行われている授業評価は, 紙もしくは電子媒体を用い, 平均して10~20個程度の項目に対して, 4~7段階 (e.g. とてもそう思う~まったくそう思わない) で評定を求めるといものである。多くの場合, 学期末に行われる総括的評価 (summative evaluation) として実施されている。

しかし, 授業評価を行う目的は単一ではなく, 重層的な機能と目的を有している。田口¹⁰⁾は授業評価の機能を以下に示す5つに要約している。

1. 意識改革の起爆剤
2. 授業改善の指針
3. 学生と教員間, あるいは教員相互のコミュニケーションツール
4. 優れた教育者を評価する, あるいは教員が自身の教育業績を示すための証拠資料
5. 教育の質保証などアカウンタビリティのための根拠資料

このうち, 多くの大学で行われている授業評価の第一義的目的は, 2.と3.に該当する授業改善である。授業評価は, 学生が授業に対してもつ多面的な認識の枠組みを明らかにすることを通して, 教員が自らの授業の“くせ”を知り, 授業実践を診断的に振り返るためのツールとして活用されている¹¹⁾¹¹⁾。その意味で, 授業評価は, 総括的評価の側面とともに, 次年度に向けた形成的評価 (formative evaluation) としての側面をも持ち合わせている。また, 授業評価は, 学生の授業に対する主観的認識を測定しているという意味で, 一種の心理尺度として

*愛媛県立医療技術大学保健科学部看護学科

の特徴を有している。そのため、後述するように授業評価は、調査研究法の観点から、多次元的授業評価尺度の有効性や、その信頼性と妥当性について検討が行われ¹⁾¹²⁾¹³⁾、さまざまな次元から各評定値 (i.e. 学生の評定値、クラス平均値) に影響を与える可能性のある要因などが明らかにされてきている¹¹⁾¹⁴⁾¹⁵⁾¹⁶⁾¹⁷⁾¹⁸⁾¹⁹⁾²⁰⁾²¹⁾。

これに対し、近年、大学教育における質保証やアカウンタビリティが強く求められる状況の中で、授業評価を一種の社会調査ととらえ、授業評価の結果を、あたかも授業を構成する各側面の実態を浮き彫りにしたものともみなそうとする立場もある²⁰⁾。この立場では、世論調査や顧客アンケートと同様に、あらかじめ“よい授業”像が想定され、その諸側面を表現した質問項目が策定される。そして、作成された評定尺度に照らして、評定対象 (i.e. 教員) が位置づけられることになる。その際、回答者 (i.e. 学生) の評価能力や多様性は問われない。その結果、各授業で設定される教育目標や教育方法の多様性は捨象され、“実態のごとくみなされた授業評価の評価点”のみが一人歩きし、評価点の高さや、場合によっては、教員間での評価点の比較にのみ関心が向けられてしまう可能性を内包している。田口¹⁰⁾の分類では4.と5.に該当するが、授業評価結果を教員評価に積極的に活用しようとする場合²⁰⁾も、この観点が色濃く反映されている。

このように、授業評価の結果に対する解釈の混乱は、授業評価をどのような性質を有する測定ツール (尺度) とみなし、そこで測定された評定値はどのような意味をもつと考えるのかという観点の違いが反映されている。

授業評価という測定ツール：質問紙調査の一つとして

さて、そもそも本稿で取りあげる授業評価は、測定論の観点からすると、質問紙尺度の一つとみなされるものである。したがって、測定 (measurement) においては、用いる尺度 (scale) の精度が重要となる。すなわち、尺度の妥当性 (validity) と信頼性 (reliability) が担保されていることが重要となる^{1)12)13)註2)}。

この点について、これまでに多変量解析等を用いて授業評価尺度の因子構造について検討した結果や、総合評価 (e.g. 授業は総合的にみて満足のいくものであった) と各設定項目との関連について検討した結果からは、“興味ある授業内容” “基本的な授業技術” “学生とのコミュニケーション” など、学生がどのような授業を望ましいと考えているかについて、ある程度一致した見解が示されている²²⁾²³⁾²⁴⁾²⁵⁾²⁶⁾。また、同一教員が担当する同一の授業科目において、受講生もほぼ毎年均質である場合、授業評価の結果としてのクラス平均値は比較的安定することが知られている⁸⁾¹⁸⁾¹⁹⁾²⁷⁾。

しかし、その一方で個々の教員の評定値は、さまざまな要因の影響を受けることも明らかにされている²⁰⁾²⁸⁾。

例えば、授業評価を実施する時期²⁸⁾²⁹⁾や教員の年齢³⁰⁾、単位の出し方¹⁶⁾²⁷⁾³¹⁾によって影響を受けることが明らかにされている。また、授業の受講者数²⁵⁾³²⁾、必修・選択の別やカリキュラムにおける位置づけ¹⁵⁾なども影響する。眞榮城・秦野³³⁾や河野・津田・寺川ほか³⁴⁾は、講義形式の授業よりも演習・実習形式の授業の方が、授業内容に対する満足度が高くなる傾向のあることを報告している。また、中井³²⁾や松本・塚本²⁵⁾は、受講者数が100名までの場合では、授業評価の結果はクラスサイズが大きくなるほど低下する傾向が見られるが、100名以上のクラスでは授業評価の評定値が上昇する傾向にあることを報告している。

一方、学生の評定行動が評定平均値に及ぼす影響についても明らかにされている。例えば、石井³⁵⁾は、いい加減な評定による端点データが4%混入すると評定平均値は0.1程度ずれることを、シミュレーションを通じて示している。5段階評定の授業評価の場合、概ね評定値のクラス平均値は3.5～5.0点で分布する。したがって、80名クラスの場合、1や5ばかりをつける極端な評定をする学生が3.2名含まれることで、評定平均値が0.1程度の変動をもたらすことになる。これは大きな影響だといわざるを得ない。

さらに、教室環境や教育設備に対する評価が低かった時には高かった授業内容への評価が、教室環境や教育設備が改善されると、かえって前年よりも下がってしまったというエピソードのように、授業評価の評定値は項目間の相対的關係の上で変動する可能性も指摘されている²⁸⁾。

このように、授業評価尺度は一定の妥当性と信頼性を有しつつも、それらに歪みを生じさせる要因もあることがわかる¹⁸⁾²⁸⁾³⁶⁾。そのため、授業評価を実施する際には、評定値に歪みをもたらす要因を極力制御するよう手続的な対策をとる、あるいは集計時に評定値の補正を行うことが求められる。しかし、そのような対策を行った上でも、授業評価がある程度の脆弱性を有することを踏まえると、その測定値が教員個人の教育活動のある側面を、実態として客観的に測定していると考えすることは危険である。むしろ、学生が認識する主観的観点を測定していると解釈することが妥当であるといえる。

にもかかわらず、時として私達は、個人の項目別や全体の評定平均値 (i.e. クラス平均値) を組織全体の平均値と単純に比較し、一喜一憂してしまうことがある。しかし、中村³⁷⁾も指摘するように、これは意味がある比較とはいえない。なぜなら、授業評価の評定平均値は、測定値という性格上、誤差を含んでおり、概ねクラス平均値±標準誤差の2倍程度の範囲に入っていれば、その差は誤差の範囲内と解釈されるべきものであるからである。同様に、個々の授業の項目別評定平均値を、受講者数や標準偏差の違いを無視して、得点の上下だけを単純

に比較することも意味がない。

したがって、各大学や教員個々人が授業評価を活用するにあたっては、分析するデータの次元（クラス平均値なのか、個々の評定値なのか）や評定値のもつ意味、授業評価の効用と限界を踏まえつつ、どのように活用していくかを考えることが重要となる。

授業評価を活用する際の理論的枠組み

これまでのところで、授業評価を社会調査あるいは実態調査とみなすことには問題があり、一つの心理尺度とみなしたうえで、その効用と限界を踏まえて活用することが重要であると指摘した。しかしながら、授業評価を学生の認知的枠組みを測定する一つの心理尺度とみなした場合でも、さらにどのような理論的枠組みを用いるかによって授業評価の活用のしかたは異なってくる。例えば、9因子からなる多次元的授業評価尺度を開発したMarshらによる先駆的研究¹¹⁾¹²⁾¹³⁾では、多様な学部・専攻から得られた授業評価データを用いて因子構造モデルの検証が行われている。その上で、単一の総合評価項目を用いることや設定された複数の評価項目の評価点を単純に縮約すること（e.g. 平均値）を避け、因子構造に基づく教員個人の授業評価プロフィールを作成することに重点がおかれている。すなわち、学生との相互作用の中で構成される授業の特徴を客観的に捉える資料を作成することが、授業評価の主目的とされている。

一方、近年注目されている立場に、顧客満足度分析（customer satisfaction analysis, 以下CS分析と略す）を応用した考え方がある²⁵⁾²⁶⁾。CS分析では、社会調査のように絶対的な指標としての“よい授業”は想定されないものの、学生を授業というサービスを楽しむ顧客とみなし（学生消費者主義：Student Consumerism）³⁸⁾、学生の授業に対する総合評価や満足度（e.g. 授業は総合的にみて満足のいくものであった）を高めることを目標とする。そして、満足度との関連から授業評価の各項目の“重要度”と“改善必要度”を算出することで、満足度に影響する教員の教授（teaching）要因を明らかにし、改善することを求める²⁵⁾²⁶⁾³⁹⁾。例えば、松本・塚本²⁵⁾は、“授業によって学習意欲や興味が増した”“授業内容はよく理解できた”“主要なポイントをはっきりと示してくれた”と学生が感じる授業ほど、総合評価（受講満足度）が高いことを明らかにしている。

もちろん、上述した二つの立場にはそれぞれ一長一短がある。例えば、Marshらのような多次元的授業評価を用いて、その結果をプロフィール化することは、教員各自に自らの“くせ”や特徴を視覚的に呈示することができる。しかし、得られた結果が意味する内容は、授業評価結果からは自明ではなく、授業科目の位置づけや授業目標、自らの実践内容を省察的にふり返ることで初めて

明らかにされる。場合によっては、評価点の高さは、必ずしも授業実践の成功を意味するとはかぎらないということも起こりうる。例えば、以下のE.P.アントナコポウロとクリス・アージリス（Argyris, C.）とのやりとりは、この点について示唆に富んでいる（中原・金井⁴⁰⁾p.125より）。

E.P.アントナコポウロ「クリス、あなたはいつもとても情熱的だが、自分の熱意や積極性が、学生にとってマイナスに作用することはありませんか？」

クリス・アージリス「……あり得ると思うよ。一部の自律的な学生にとっては、教える私が情熱的であるがゆえに、かえって学習が阻害されるのではないかと考えられる。」

また、結果を授業改善につなげる場合でも、実際に何をどのように改善すればよいのかについて、授業評価自体は明らかにしてくれない。授業評価結果は授業をふり返る資料にすぎず、改善の必要の有無や改善策は、学生との相互作用による結果として自己省察的に教員が判断する、あるいは、授業評価結果を基にした教員相互の議論や授業のピア・レビューを通じて判断されることになる。このことは、教員と学生あるいは教員相互のコミュニケーションツールとしての機能¹⁰⁾が、授業改善やカリキュラムを含めた教育改善にとって重要となることを示している。

これに対し、CS分析の考え方は、各教員が何を目標として授業改善を行えばよいのかについて一定の方向づけがなされる点や、優先すべき改善点を明確にできるという点で²⁰⁾²⁵⁾²⁶⁾、授業改善に向けて取り組みやすい第一歩といえる³⁹⁾。しかし、授業評価を一種の顧客アンケートとみなす場合、教員の提供する授業という教育サービスについて、学生が満足度を中心とした既存の価値観や感覚に基づいて評定することを許容することとなる。確かに学生の受講満足度は、授業改善にとって重要な指標の一つではある。しかしながら、CS分析では授業者（教員）と学習者（学生）との相互作用や、学習過程で生じる学習者の価値観の再構成といったダイナミクスの観点（e.g. どのような学習経験の結果として、受講満足度が導かれているのかといった点）は重視されていない⁴¹⁾。また、授業評価の結果を個人間で比較することを許容し、教員間で“上手な授業”を目指し競い合う傾向を助長する可能性も内包している²¹⁾。その一方で、各項目の評定値が既にはほぼ天井効果を示している（e.g. 5段階評定で評定平均値が4.0以上を示す）教員にとっては、不必要な教授技術の改善努力を継続的に強いられることになり、疲弊感のみを増長させる可能性もある。

このように、両者の立場には、それぞれ一長一短があることがわかる。しかし、CS分析の観点では、評定値を価値づけ、授業の成否を評価するのは、本来アンケートの回答者である学生ではなく、授業の企画・運営者と

しての教員であるという観点が欠如しているという点は、後述する教育学的観点からは大きな問題点といえる。

“教授・学習”の観点、学生の心理的特性との関連

前節で紹介した点に加え、教育学的観点からは授業評価にはもう一つの大きな課題がある。それは、多くの場合、授業評価は授業時間を割いて行われるにもかかわらず、もっぱら“学生が授業をどうみたか”“学生が授業に満足したか”を尋ねる項目に比重が置かれる傾向があり、学生自身が“授業で何を学び、どのような力をつけたか”を自己省察的に振り返る学習としての機会は、ほとんど与えられていないという点である^{38)42)43)44)45)註3)}。この点は、授業評価を教育評価という観点からとらえた場合、重要な指摘といえる³⁸⁾。すなわち、授業評価を教育活動の一部ととらえ、定期試験による結果だけではなく、自己省察 (reflection) を通して、学生に自身の学びのプロセスをメタ的に評価させるという教育的観点が、今日の授業評価の考え方には欠如していることを意味している。そのため、授業という学習空間では、“教員による授業の準備や工夫だけでなく、教員と学生あるいは学生同士の相互作用が、学生の学びや理解を促進し⁴²⁾⁴⁴⁾、その結果として受講満足度が高くなる”というグループダイナミクスのプロセスが仮定される³⁸⁾¹⁹⁾⁴¹⁾、CS分析をはじめ現在の取り組みでは、“教員の教授 (teaching)”と“学生の学習 (learning)”“受講満足度”の3つの視点を考慮して授業評価が作成されている実践例は極めて少ない⁸⁾⁹⁾⁴⁶⁾。

その中で、澤田⁹⁾は、大塚¹⁸⁾や松下³⁸⁾などの先行研究を基に、教員の教授面と学生の学習面の両側面について評定を求める多面的授業評価尺度を作成している。そして、教員が教授技術を改善することにより、たしかに学生の受講満足度を高めることは可能であること、しかし、学生の学びを考えると、学習コミュニティ (learning community)¹⁸⁾への参加を促す取り組みをすることで、学生自身が自らの学びの有意義性を認識し、その結果として“受講満足度”を高めていけるような工夫を行うことが重要であることを示している。澤田⁹⁾の指摘は、教員の教授技術の改善と学生の満足度からのみ授業評価をとらえる (e.g. CS分析) のではなく、教授と学習の両側面から結果をプロフィール化し、教員の具体的な授業実践 (e.g. 学習コミュニティの活用) に照らした仮説検証作業として、授業評価を活用することの必要性和有効性を示すものといえる。そうすることにより、学生が授業に不満を感じているか否かを把握するとともに、教員が設定した授業目標 (e.g. 学生相互に学び合う、批判的思考ができるようになる) を、学生が学習成果として達成できているか否かを把握することができる。また、授業評価のクラス評価点は高いにもかかわらず、定期試験

で期待された教育成果が得られないといった、教員が経験しやすい状況にも対処しやすくなると考えられる。

さらに、もう一つの課題として、従来、授業評価の評定値は、学生の心理特性と切り離してとらえられてきたという点が挙げられる。この点についても、近年では、むしろ教員と学生との相互作用や学生の心理特性と関連づけて、授業評価を検討しようとする動きもみられる⁸⁾⁹⁾³⁶⁾⁴⁷⁾⁴⁸⁾。その結果、授業評価の結果は、教員と学生との相互作用の文脈に規定されることが明らかにされてきている。

例えば、豊田・中村⁴⁷⁾は、教授者、学生、項目(観点)、順序の4つの要因と交互作用要因による評定値の変動の大きさを定量的に検討することで、授業評価の信頼性について検討している。その結果、最も評定値の変動への寄与の大きい要因として“教授者と学生との交互作用要因”を、次に大きい要因として“学生の要因”を検出している。

このうち、前者の知見については、いわゆる“Dr. Fox効果”を検討したMarsh & Ware, Jr.¹⁷⁾も、学習動機づけの高い学生の場合では、教員のパフォーマンスは授業評価に影響せず、授業内容の質 (体系性、情報量) が影響するのに対し、学習動機づけの低い学生の場合では、教員の授業パフォーマンスが授業評価結果に影響することを明らかにしている。一方、後者の知見については、田中・藤田⁴⁸⁾や澤田⁸⁾⁹⁾が、学生の心理的特性として達成目標 (achievement goals) を取りあげ、授業評価の諸側面との関連を検討している。ここで達成目標とは、学習動機づけの質的側面を指しており、ある達成場面でも何を成功あるいは失敗とみなすかなどの認知的な枠組みとなるものである。田中・藤田⁴⁸⁾は、達成目標を修得 (mastery) 志向、パフォーマンス接近 (performance-approach) 志向、パフォーマンス回避 (performance-avoidance) 志向の三つの目標に分類して授業評価との関連を検討した^{36)48)註4)}。その結果、他者との相対比較によるのではなく、自分自身が学習や理解を通じて自己の能力を高めたいとする修得志向の高い者ほど、授業に対する肯定的評価が高いことを明らかにしている。

これらの研究は、授業評価の評定値の高低のみならず、どのような特性を持つ学生により高い評価が得られているのかを検討することの必要性を提起しているといえる。

まとめ：授業評価の有効活用に向けて

授業評価は教員集団が教育活動をふり返り、授業改善につなげるための方法の一つにすぎない。しかし、公開授業や種々の研修会と比べても、最も研究的視点からの検証が進んでいる授業改善ツールである。本稿で考察したように、授業評価は、教員が自律的に自分自身の授業を振り返り、改善を志向する形成的評価 (formative evaluation) としての資料として用いるのであれば、き

わめて有効な情報源となる。そのためには、学期末に一種の義務として受動的に授業評価を行うのではなく、自らの授業実践について仮説検証を行うという視点をもって授業評価を分析・活用することが望ましいといえる。例えば、ミニツツペーパーやコミュニケーションカードの活用など、日頃の形成的評価を組み合わせ、結果を分析することも一つの方法といえる。

その上で今後、評価活動の外圧に振り回されず、教員個々あるいはFD活動として授業評価を有効に活用していくための方向性として、本稿では以下の2点について提言することができよう。

第1に、授業評価がもつ効用と限界を踏まえた上で、教員の授業技術や学生の満足感のみに着目するのではなく、授業評価を教育活動の一環ととらえ、教授と学習 (teaching and learning) の両側面を測定することができる多面的授業評価を活用していくことの有効性についてである。このことは、以下の点で教員と学生の両者にとって有用であると考えられる。まず、学生にとっては、自身の学びについてリフレクティブに評定することで、当該授業の学習目標を想起し、なによりコースを通じての自己の学びの成果や意味、不十分さをメタ認知的に振り返る貴重な学習体験の場となる。一方、教員にとっても二つの点で有効性をもつ。基本的な教授スキルの面で評価点が低い者 (e.g. 5段階評価で3.5点未満) にとっては、学生の学習成果との関連において自己の改善点が明確になり、従来以上に具体的な改善行動につなげるプレッシャーとなる。これに対し、教授面での評価点が一定以上高い者にとっては、成績評価と学生の自己省察というアウトカムの両側面から自らの授業実践を振り返り、教授スキルにとらわれず、教材内容や教育方法などを検討するための資料とすることができると考えられる。

第2に、豊田・中村⁴⁷⁾や田中・藤田⁴⁸⁾が示すように、授業評価の評定値が教員と学生の相互作用に規定されることを考慮すると、授業評価の評定値を学生のもつ心理的特性と切り離してとらえるのではなく、どのような特性をもつ学生が多面的な授業評価のどのような側面 (e.g. 授業内グループダイナミクス、学習成果の自己評価、満足度) を高く評価する傾向にあるのかを検討することの有効性についてである。そのことにより、自分自身の教育実践が、学生との相互行為を通して、授業目標のどのような点が達成でき、どのような点が達成できなかったのか、といった点について、自己省察および相互研修型FDの場を通じて仮説検証していくことができると考えられる。

最後に、現在多くの大学では、授業評価は匿名式でおこなわれている。そのため、授業評価で測定される各側面 (e.g. 教員のteaching, 学生のlearning, 受講満足度)

と学生の学業成績との関連については、十分に検討されているとはいえない³⁶⁾³⁸⁾。僅かに藤田・溝上²⁹⁾や田中・藤田⁴⁸⁾、松下³⁸⁾などで検討されているにすぎない。しかし今後は、学生の学習成果と授業評価を関連づけて検討することも必要となると考えられる。

文 献

- 1) Marsh, H. W. (1987) : Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, 11, 253-338.
- 2) 読売新聞教育取材班 (2009) : 「教育ルネッサンスー大学の実力ー」. 中央公論新社
- 3) 高橋和子, 林義樹, 種田保穂ほか (2005) : 授業改善に向けた全学の取り組みー授業評価と授業改善計画書の一体化ー. *京都大学高等教育研究*, 11, 19-32.
- 4) 佐藤浩章 (2008) : 愛媛大学におけるFDの取り組み. 「研究・教育のシナジーとFDの将来」. 東北大学高等教育開発推進センター編, p.83-99, 東北大学出版会
- 5) 佐藤浩章 (2009) : FD における臨床研究の必要性とその課題ー授業コンサルテーションの効果測定を事例にー. *名古屋高等教育研究*, 9, 179-198.
- 6) 飯島優雅, 深尾暁子 (2008) : 国際基督教大学における授業評価能力養成の試み. *大学教育学会誌*, 30, 118-122.
- 7) 佐藤手織, 松浦勉, 小林繁吉ほか (2009) : 授業評価に影響を与える要因ー特に「信頼性」をめぐる考え方についてー. *八戸工業大学異分野融合研究所紀要*, 8, 61-78.
- 8) 澤田忠幸 (2007) : 学生による授業評価と心理的特性との関連ー達成目標および特性的自己効力感を用いてー. *京都大学高等教育研究*, 13, 1-12.
- 9) 澤田忠幸 (2009) : 多次元的な“学生による授業評価”活用の有効性ー学習コミュニティおよび学習成果の自己評価の観点を含めてー. *大学教育学会誌*, 31, 132-139.
- 10) 田口真奈 (2007) : 授業評価の諸機能. 「授業評価活用ハンドブック」. 山地弘起編, p.31-51, 玉川大学出版部
- 11) Marsh, H. W., & Roche, L. A. (1997) : Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias, and utility. *American Psychologist*, 52, 1187-1197.
- 12) Marsh, H. W. (1991) : Multidimensional students' evaluations of teaching effectiveness: A test of alternative higher-order structures. *Journal of Educational Psychology*, 83, 285-296.

- 13) Marsh, H.W. & Hocevar, D. (1991) : The multidimensionality of students' evaluations of teaching effectiveness : The generality of factor structures across academic discipline, instructor level, and course level. *Teaching and Teacher Education*, 7, 9-18.
- 14) Greenward, A.G. (1997) : Validity concerns and usefulness of student ratings of instruction. *American Psychologist*, 52, 1182-1186.
- 15) 林田 実, 藤井 敦 (2005) : 授業評価決定因子の統計的研究. *大学教育学会誌*, 27, 82-92.
- 16) Marsh, H.W., & Roche, L.A. (2000) : Effects of grading leniency and low workload on students' evaluations of teaching : Popular myth, bias, validity, or innocent bystanders. *Journal of Educational Psychology*, 92, 202-228.
- 17) Marsh, H.W., & Ware, Jr, J.E. (1982) : Effects of expressiveness, content coverage, and incentive on multidimensional student rating scale: New interpretations of the Dr. Fox effect. *Journal of Educational Psychology*, 74, 126-134.
- 18) 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. 「心理学者, 大学教育への挑戦」. 溝上慎一・藤田哲也編, p.1-38, ナカニシヤ出版
- 19) 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, 102-109.
- 20) 安岡高志 (2007) : 学生による授業評価の進展を探る. *京都大学高等教育研究*, 13, 73-87.
- 21) Wachtel, H.K. (1998) : Student evaluation of college teaching effectiveness : A brief review. *Assessment and Evaluation in Higher Education*, 23, 191-211.
- 22) 八木紀一郎 (2004) : 学生による授業評価データの分析と解釈. *京都大学高等教育研究*, 10, 59-66.
- 23) 守 一雄, 野口宗雄, 筒井健雄ほか (1996) : 学生の授業評価による望ましい大学授業の特質の解明(2) - 総合評価を目的変数とする重回帰分析結果 -. *信州大学教育学部紀要*, 89, 65-73.
- 24) 安岡高志, 峯崎俊哉, 高野二郎ほか (2002) : 授業アンケートにおける学生の達成感と総合評価の関係. *大学教育学会誌*, 24, 123-126.
- 25) 松本幸正, 塚本弥八郎 (2003) : 学生による授業評価アンケート結果の分析手法 - 授業改善の定量的評価手法の確立を目指して -. *名城大学理工学部研究報告*, 43, 80-87.
- 26) 松本幸正, 塚本弥八郎 (2004) : CS分析の考え方を導入した授業評価アンケートの分析と授業改善ポイントの定量化. *京都大学高等教育研究*, 10, 21-32.
- 27) 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, 68-74.
- 28) 大山泰宏 (2001) : 大学教育評価の課題と展望. *京都大学高等教育研究*, 7, 37-55.
- 29) 藤田哲也, 溝上慎一 (2001) : 授業通信による学生との相互作用 I - 学生はいかに「藤のたより」を受け止めているか -. *京都大学高等教育研究*, 7, 71-87.
- 30) 安岡高志, 峯崎俊哉, 山本銀次ほか (1997) : 学生の授業評価におよぼす教員の年齢の影響. *大学教育学会誌*, 19, 75-79.
- 31) 南 学 (2006) : 単位認定が「甘い」授業の評価は「甘く」なるのか - 成績判定指標と学生による授業評価の関係 -. *松山大学論集*, 17, 141-155.
- 32) 中井俊樹 (2006) : クラス規模は授業にどのような影響を与えるのか. *名古屋高等教育研究*, 6, 5-19.
- 33) 眞榮城和美, 秦野悦子 (2004) : 学生による授業評価(2) - 2004年度前期調査 -. *白百合女子大学研究所紀要*, 40, 109-124.
- 34) 河野和明, 津田早苗, 寺川みち子ほか (2004) : 学生による授業評価の方法論的考察 - 東海学園大学人文学部における2002年度秋学期の結果から -. *東海学園大学研究紀要*, 9, 117-139.
- 35) 石井秀宗 (2006) : いい加減な評定および打ち切りに関する一検討. *日本心理学会第70回大会ワークショップ 心理学者, 授業評価への挑戦 2 - 測定・評価・活用 - 配付資料*.
- 36) 藤田哲也 (2005) : 授業評価に対する心理学的アプローチ. *名古屋高等教育研究*, 5, 267-280.
- 37) 中村知靖 (2007) : 授業評価アンケートの整理. 「授業評価活用ハンドブック」. 山地弘起編, p.122-138, 玉川大学出版部
- 38) 松下佳代 (2005) : 学生による授業評価 - 現状と課題 -. *京都大学高等教育叢書*, 21, 203-208.
- 39) 南 学 (2007) : 学生による授業評価へのCS分析の適用. *三重大学教育実践総合センター紀要*, 27, 29-34.
- 40) 中原 淳, 金井壽宏 (2009) : 「リフレクティブ・マネジャー - 一流はつねに内省する -」, 光文社
- 41) 大山泰宏 (2007) : 授業評価の発想と歴史. 「授業評価活用ハンドブック」. 山地弘起編, p.11-30, 玉川大学出版部
- 42) Cashin, W.E., & Downey, R.G. (1992) : Using global student rating items for summative evaluation. *Journal of Educational Psychology*, 84, 563-572.
- 43) Mckeachie, W.J. (1997) : Student ratings : The validity of use. *American Psychologist*, 52, 1218-1225.
- 44) Sheehan, E.P., & DuPrey, T. (1999) : Student evaluations of university teaching: Statistical data included. *Journal of Instructional Psychology*, 26, 188-193.
- 45) 串本 剛 (2005) : 教育目的との対応にみる教育評

価の妥当性－授業評価項目の分析を具体例に－. 大学教育学会誌, 27, 124-130.

- 46) 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, 463-473.
- 47) 豊田秀樹, 中村健太郎 (2004): 大学における授業評価の信頼性－一般化可能性モデルと構造方程式モデルによる4相データの解析－. 心理学研究, 75, 109-117.
- 48) 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, 397-403.
- 49) 大塚雄作 (2007): 授業評価とアカウンタビリティ. 「授業評価活用ハンドブック」. 山地弘起編, p.80-101, 玉川大学出版部

註

- 1) 大塚⁴⁹⁾によると, 授業評価の普及は, 大学評価に関する施策と対応している。その契機は, 1991年の大学審議会答申「大学教育の改善について(答申)」による自己点検評価の努力義務化と, それに伴う大学設置基準の「大綱化」に求めることができる。その後, 「21世紀の大学像と今後の改革方策について(答申)－競争的環境の中で個性輝く大学－」(1998)において, 第三者評価システムの導入が提言されたことを受け, 翌1999年にはFD実施の努力義務化(大学設置基準)および自己点検評価の義務化が示され, 続く2000年には大学評価・学位授与機構による試行的第三者評価の導入等が行われた。2004年度からは認証評価機関による機関別認証評価制度が発足している。この間に, 1992年には10%にも満たなかった授業評価の実施率も, 1998年には50%を越え, 2004年には97.5%に達している⁴⁹⁾。さらに, 2006年の改正教育基本法では教員の資質向上が明記され, 2008年度より学士課程におけるFDの義務化が謳われている。今日に至っては, ほぼ全ての大学で何らかの形で授業評価は行われており²⁾, 今や授業評価導入の是非を検討する段階から, 自律的に有効活用のあり方を検討する段階に入っているといえる。
- 2) 妥当性 (validity) とは, 用いる尺度が本来測るべきものを正確に測れているかどうかを指している。その基準としては, 内容的妥当性や構成概念妥当性, 基準関連妥当性などがある。一方, 信頼性 (reliability) とは, 用いた尺度により測定された結果が安定性や一貫性を有しているかどうかを指している。再テスト法やクロンバックの α 係数など内的整合性により検証される。一般的な表現としての回

答者の信頼性とは区別されるものである。

- 3) 串本⁴⁵⁾は, 我が国で使われている100種類の授業評価票を調査し, 授業評価項目の中には学生の取組や受講態度を測定する項目が2割以上含まれていたことを報告している。しかし, これらの項目は, 主として授業評価結果の分析の際の制御変数としての役割や学生の真摯な回答態度への方向付けを意図したものが多く, ここで指摘する学生が自己の学び (learning) の側面を振り返る側面とは異なると解釈される。
- 4) 田中・藤田⁴⁸⁾は, 達成目標を三つに分類している。一つは, 他者との比較ではなく, 自分自身が学習や理解を通じて能力を高めたいとする“修得 (mastery) 志向”の目標である。もう一つは, 他者に対してできるだけ少ない努力で相対的に優位になることで能力の高さを誇示したいとする“遂行志向”の目標である。後者はさらに自分の有能さを誇示しポジティブな評価を得ようとする“遂行接近 (performance-approach) 志向”と, 自分の無能さが明らかになる事態を避け, ネガティブな評価を回避しようとする“遂行回避 (performance-avoidance) 志向”に分類される。また, 澤田⁸⁾は, 日常行動全般に対する自己効力感 (generalized self-efficacy : 一般的自己効力感) の高い人ほど“修得志向”の達成目標が高く, “遂行回避志向”の達成目標が低いことを明らかにしている。

要 旨

今日, 学生による授業評価アンケートを実施している大学は, 国公私立を問わずほぼ100%に達している。また, 授業評価を有効に活用するための様々な組織的取り組みも行われ始めている。しかし, その一方で, 実施の形骸化や効果への疑問の声も囁かれ始めている。その背景には, かつてのような授業評価尺度の信頼性や妥当性, あるいは学生の評定能力に対する疑念や不信とは異なり, 授業評価の結果をどのように解釈し, どのような理論的枠組みに基づいて授業評価を活用すればよいのかに関する戸惑いや混乱がある。

本稿では, 授業評価を活用する理論的枠組みを整理するとともに, 評価測定論および教育学的観点から, 授業評価研究が明らかにしてきた知見を振り返った。その上で, 授業評価の効用と限界を踏まえつつ, 有効に活用していく方向性の一つとして, 教授・学習双方の側面について測定が可能な多面的授業評価の活用や, 学生の特性要因と関連づけた分析の有効性について考察した。