

データ収集による高等教育分野における 教育アセスメントと評価への貢献

有本昌弘*

西塚孝平**

本学部研究科での授業評価は、これまで10年以上にわたって続けてきているが、教員にとって、評価はしなければならないものという考えのもとルーチンで行われ、学生とアセスメント自体を楽しむという発想はなく、目立った成果は挙げられていない。この辺りは、海外と大きく乖離が見られる。ここにきて、本研究科もオンライン中心の講義・演習となったことに伴い、学生からの講義変革や学びのイノベーションへの要望は強くなり、教員側の意識変革もある種のモーメントを迎えている。その中で、混合研究法のデータ収集により日本のアセスメント研究の存在感を示した Wicking (2020) の補足と解説を吟味検討することは、高等教育分野におけるアセスメント観を転換させるための起爆剤になりうる。当該論文の課題を挙げて展望を示すことにより、さらなる国際共同研究を推進する礎を築き、また、FD としても応用可能なものになっていくだろう。

キーワード：形成的アセスメント、日本の高等教育、学習志向のアセスメント、儒教的伝承文化 (CHC)、語学教育

1. はじめに

一般のオンライン授業により、また LMS (学習管理システム) の進化により、格段にデータ入手と整理分析が容易となり、にわかに、フィードバックなどを中核にしたアセスメントが注目されつつある。そこで、標記のテーマに迫るため、1975年に vol.1 を出し45年継続されている *Assessment & Evaluation in Higher Education* をレビューしたが、日本の高等教育実践のデータを使って論じたものは皆無に近い状況であった。和文はおろか英文での当該分野での発信は、特に国内の教育学部・教育学研究科、他大学からは皆無に近い。そうした中で、最新号に掲載された日本のウィキンによる論文 [Wicking, P. (2020) . Formative assessment of students from a Confucian heritage culture: Insights from Japan. *Assessment & Evaluation in Higher Education*, 45:2, 180-192.] は例外であり、混合研究法のデータ収集によって形成的アセスメントの存在感を示したものである。

そこで、当該論文の著者のご厚意により、Wicking (2020) の全訳と補足と解説、吟味検討を行うとともに、課題を挙げながらさらなる国際共同研究を推進する礎として提案する。

*教育学研究科 教授

**教育学研究科 博士課程後期/日本学術振興会特別研究員 DC1

2. Wicking (2020) の本文全訳

抄録

今世紀に入ってから教育アセスメントへの関心はますます高まっており、学習プロセスの成果を向上させたり、逆に低下させたりするともいわれている。とりわけ脚光を浴びているのは形成的アセスメントと呼ばれる、活動をアセスメントすることによって学習を促進させる実践である。しかし、儒教的伝承文化の社会においては、総括的アセスメントのほうが重要度は高いものと広く信じられており、形成的な側面を蔑ろにしている。その社会にいる学習者がアセスメント課題に挑戦するとき、課題の捉え方や取り組み方は儒教的伝承文化の影響を受けているにちがいない。そこで本研究では、儒教的伝承文化にいる学習者の気質を理解するために、質的・量的データの分析を試みる。そして、日本の大学生が所有しているアセスメントの概念と経験に注目し、形成的アセスメント実践を今後発展させていくにあたっての方向性を提示する。分析データの収集には、自己申告調査法とナラティブの文章完成法の二つを活用した。その結果、日本の大学生は典型的な儒教的カテゴリーに当てはまらない気質も発揮しながら、アセスメント課題に取り組んでいることが分かった。そのカテゴリーとは、過剰な競争意識、家族を喜ばせる心、実践スキルよりも座学で得た知識の重視、などである。考察の結論部では以上を踏まえ、教育者に向けたペダゴジカルな展望を指摘した。

社会的背景^(注1)

儒教的な考えに基づく中国の科挙制度は、世界中の教育アセスメントに多大な影響を与えてきた。今では標準化された大規模テストが盛んに実施されているが、その起源はこの科挙制度にまで遡ることができる。当時は暗記・詰め込み学習が重要視されたため、科挙制度の主眼は受検者の知識を総括的に価値判断することに置かれていた。今日になってようやく総括的テストとは距離を取り始めた代わりに、形成的アセスメントの有効性が広く知られるようになってきた。しかし、儒教的伝承文化(CHC)を共有する諸国では、この制度の名残がまだ根強くある。CHCにいる学習者が英語圏に留学するとしたら、アセスメントに対して極めて異質な捉え方を思いがけず経験するのではないだろうか。したがって、学習者に応じてアセスメントの印象や経験が異なるのかどうかを理解できれば、学習の成長を阻害するのではなく促進させるアセスメント実践を教育者が導入できるようになるだろう。

これまで、形成的な目的を果たすアセスメントが言語学習の利となることは広く認められてきた。アセスメントは学習者を能力別に選抜、順位づけするだけではない。アセスメントの形成的な捉え方とは、目の前の学習を具体化し、未来の目標を明確化しようとするものである。同時に学習者は、目標を手繰り寄せるための豊かな学習プロセスをも内的に発達させていく。このとき形成的アセスメントは、学習者の感情や学習意欲、信念、態度、概念といった学習とアセスメントに関わるものに触れながら、学習の情動的・認知的な次元の中で理解せざるを得なくなる。

アセスメントの知覚は、学習に向かう姿勢に対して大いに影響を及ぼしている (Struyven, Dochy

& Janssens, 2005)。その知覚は自身の隠れたカリキュラムを構成し (Sambell & McDowell, 1998), 学習者の経験と信念に基づく独自の意味をアセスメント課題に付与する。学習者は、仮に明示されていたとしても、アセスメントの目的を理解できない場合が多いだろうし (Rea-Dickins, 2006), 付け加えるなら、課題に込められた目標に毅然と逆らって、自らの目的を達成しようとすることさえあるだろう (Spence-Brown, 2001)。学習をどの程度行うのかの大部分が学習者によって決まってしまうことは、形成的アセスメント実践を検討するにあたって重要な事実である (Davison & Leung, 2009)。そこで本研究では、日本という一つのCHCに生きる学習者に着目し、どうすれば形成的アセスメントを上手く導入できるようになるのかについて、理解を深めていくこととする。

形成的アセスメントの理論的枠組みは、主に英語圏の文脈での研究を中心に数多く提案されてきた。それらは、真正のアセスメント (Finch, 2002; Frey et al., 2012), ダイナミック・アセスメント (Poehner, 2007; James, 2012), 学習のためのアセスメント (Assessment Reform Group, 2002; Gardner, 2012; Sambell, McDowell & Montgomery, 2013), 教育者を基軸にしたアセスメント (Davison & Leung, 2009), 学習志向のアセスメント (Carless, 2011; Jones & Saville, 2016; Turner & Purpura, 2016) などと称されている。個々に違いはあれど、これらは次の三つの原則に共通性があり、アセスメントが学習を支えるものと考えられている。第1に、アセスメント課題は学習志向的であり、仲間や教育者とともに社会的相互作用の中で行われることである。多彩な課題に取り組むことで、様々な角度から達成度を価値判断することができる。第2に、フォーマルとインフォーマルな情報源に基づく、具体性に富んだフィードバックに学習者を関与させようとすることである。第3に、自分(や他者)のパフォーマンスを価値判断できるような専門性を発達させることである。そのために、ピア・アセスメントやセルフ・アセスメントといった方法や、ループリックや学習者自身が開発した採点のクライテリアなどの道具を使用することが考えられる。世界的にみても、これらの原則の多くは多国籍から成る大学教室に広く取り入れられている。しかし、そうした実践が珍しい文化の下で学習者が育つ場合、彼らに対して形成的アセスメント実践を導入するとなれば問題が生じるかもしれない。

当然のことながら、学習者がアセスメントをどのように知覚し、経験するのには個々に様々である。それは共有された文化的伝承とは全く関係がないこともある。例えば、彼らの学習履歴、学校教育環境、家庭教育、社会経済的背景なども要因に入ってくる。しかし一方では、ある広範な文化的傾向が学習者の気質を決定づけうることも分かっている。これまで多くの東アジア諸国家は儒教的伝承文化というカテゴリーに分類されてきた (Hofstede & Bond, 1988)。これには日本、香港、中国、台湾、シンガポール、韓国などが該当する。以下四つの主原則が儒教的指導には内包されており、その文化を下支えしている。(1)人間同士の不平等な関係性によって社会は安定している、(2)社会組織の基本単位は家族である、(3)他者への徳のある行動とは、自分がそうしてほしいと思うことを相手にすることである、(4)人生における徳のある行動とは、スキルや学習の修得に励み、懸命に努力し、必要以上にお金をかけず、根性をもって忍耐強くいることである (Hofstede & Bond, 1988)。

この諸原則は、儒教文化の下で成長する人々の世界観を支えているとされており、教育やアセスメントも含めて、経験を意味づける方法に働きかけている。具体的に教育を取り上げたときに、ハンとヨウ (Han & Yang, 2001) は次の四つの観点に教育への儒教的志向性が残存していると指摘する。第1に、何よりも教育は実用性にその本質がある。第2に、試験には重要な役割があるため、それを上手くやり遂げることは目の前の学習よりも重要な意味をもっている。第3に、座学で獲得した知識を実践的スキルよりも優先する。第4に、形成的アセスメントをほぼ無視して総括的アセスメントを好む。このように考えると、この4観点は形成的アセスメントの原則とは真逆のように思われる。その原則は、試験ではなくその代わりとなるアセスメント課題を重んじるとともに、教科書で学んだ知識ではなく実践的スキルの実質的な成長を促そうとする^(注2)ものだからである。

カーレス (Carless, 2011) によれば、CHCの世界観が実に多くの場面でアセスメント実践を方向づけている。例えば日本においては、ハイスティクスなテストに備えるときの苦しみに耐え抜くことそれ自体に美德があると考えられている。日本の学習者が褒められるのは、心身の重圧に耐えきって入試に備えたことに対してである (Kerr, 2001)。そして、このプロセスの中で、忍耐力や自制力の価値を学習していく。もはやハイスティクスな試験は人間の知識を価値判断するために機能してはならず、そこで測定しているものは学習者の忍耐力、決断力、そして、教育システムの頂点に立つ権威への忠誠心だという主張もある (McVeigh, 2006)。さらにそこでは、競争を極めて重視することも考えられている。学習者は相対的に順位づけされるが、そのうち上位層にはごくわずかな人数しか入ることができないため、互いに競争し合いながら高みを目指すのである。おそらく、競争が努力を喚起することで高水準の達成をもたらし、社会的利益や経済発展へとつながるのだろう。このように、成績のために競争する文化では、勤勉、忍耐、自制のできる人が恩恵を受けており、それは能力主義システムと呼ばれうるものに包摂されている (Hood, 2001)。

CHCの教育アセスメントに対する批判には、暗記・詰め込み学習の重視は実質的に無益であること (Kerr, 2001)、また、事実に基づく推論や論拠に基づく推理、直感的な推測の能力が貧弱になってしまうといった指摘がある (McVeigh, 2002)。しかしながら、そうした主張を支持するエビデンスは十分ではない。一般的にいえば、韓国や日本などの諸国に対する教育批判の多くは、「標準化された教育がもつ負の影響に関しては、未検証で、極端に単純化された、先入観付きの仮定に基づいている」(Park, 2013, p.3) ように思われる。むしろかえって、欧米諸国では暗記学習として冷笑される学習のほうが、実際には深い学習や認知的関与がみられうるとする主張もあるほどである (Biggs, 1999; Hofstede, Hofstede & Minkov, 2010)。ある事例では、日本と台湾の大学入試問題を内容分析したところ、宣言的知識が極めて重要ではあったが、それらは分析的で解釈的な翻訳スキルを強く要求する問題でもあったという (Baumgart & Halse, 1999)。したがって、CHCは未だ論争的な概念であり、様々な理解と解釈の葛藤が、諸国の間や内部で起きている。

現在までに、儒教的背景をもった社会にいる学習者に対して、どうすれば形成性を目的にしたアセスメントを駆動させることができるだろうか、という問いはほとんど解明されていない。このとき、アセスメントへの理解を深めずに、英語圏の文脈由来の研究を過信するだけでは、形成的実践

を広めんとするアセスメント改革の挑戦は袋小路に入ってしまうだろう。それはまさしく、香港の二の舞である。香港政府は形成的アセスメントの理念を導入しようと全力を注いだが、社会に深く根づいた試験文化や社会的価値観、経済的要求が、この教育イニシアティブを阻むものとなった(Berry, 2011)。韓国においても、数多くのパッチワーク的な改革は成功せず、ハイスティクスな大学入試が社会に及ぼしている影響を断ち切ることはできなかった(Kwon, Lee & Shin, 2015)。より広範な社会文化的・思想的環境が、アセスメントの信念と実践に強く働きかけている(Teasdale & Leung, 2000)。アセスメント実践の改善に向けて道案内してくれる研究を数多く築き上げるには、多背景にわたる参加者を対象にした多くの研究が必要とされる(Hayward, 2012; Reimann & Sadler, 2017)。

本研究で焦点を当てるのは高等教育の学生であり、未だ研究の盲点となっている領域である(Lopez-Pastor & Sicilia-Camacho, 2015)。ここでは特に、CHCの教育システムにみられる五つの特徴を扱っていく。この特徴は先行研究から得られたものであり、キャリアと社会進出、家族的恩義、競争、座学で得た知識か実践力か、そして実用性である。

課題設定

本研究の主な関心は次の二つの問いである。一つは、CHCにいる学生はアセスメントをどのように捉え、経験しているのだろうか、である。この問いは、形成的と総括的な目的を双方含んでいるアセスメント一般を取り扱う。もう一つは、この問いの答えを踏まえ、学習者にとって益となりうる形成的アセスメントを教育者がどのように実装できるのだろうか、である。学習者のアセスメント概念は学力到達度に作用しているため(Brown & Hirschfeld, 2008)、その概念を理解できれば、教育者は学習成果をより高めていく形でアセスメントに着手できるようになるだろう。国籍混成型の授業では、アセスメント概念を自覚することでアセスメント課題を創り出すことが期待できる。それは以前よりも平等性が担保され、単一の文化的集団を特別扱いすることのないものである。学習者がどのようにアセスメントを経験するのかがみえてくると、様々なアセスメント方法が学習の気質や意欲にどのような効果を出しうるのかもまた解明されるだろう。本研究では、分析データを量的と質的な情報源から取得し、統合させていく。その結果、いくつかの共通した傾向とテーマが学生のアセスメント概念と経験から現れていることが示された。それは、政策と実践への提言を行うにあたっての基盤となるものである。

方法

日本の大学でのデータ収集は次の二つの方法を用いた。一つ目は学習者のアセスメント概念リスト(SCoA)として知られている、多元的な自己申告調査法である。この方法はニュージーランドの中等教育の生徒に初めて使用された(Brown & Hirschfeld, 2008)。その後、SCoAは多くの改良が加えられており、今日までにブラジル、アメリカ、オーストラリア、ニュージーランド、そしてイランの高等教育で活用されている(Brown, Pishghadam & Sadafian, 2014)。つい最近、中国と香港の

高等教育の文脈に適した改訂版が作成された (Brown & Wang, 2016)。それは学習者のアセスメント概念リストの中国版 (C-SCoA) と呼ばれるもので、日本では一度も使われたことがない。C-SCoA は妥当性を検証する厳格な試験と予備調査を事前にクリアしており、中国の文脈に相応しいことが証明されている。その調査が設計された本来の目的は、学習者がアセスメントの用途と効果をどのように考えているのかを知ることにあつた。C-SCoA は32の項目から構成されており^(注3)、1(そう思わない)から6(そう思う)の6件法で回答を求める。元の C-SCoA は多様な構成概念を測定するために設計されているが、本研究では、その中から五つのみを取り出した。それは、キャリアと社会進出、家族的恩義、競争、座学で得た知識か実践力か、そして実用性である。

二つ目のデータ収集方法はナラティブ記述である。ナラティブ的探究は教育学研究においてここ20年で発展してきた。ナラティブ研究は現象の複雑な全体像を描き出すため、問題を多角的に認識し、研究の新しい道を切り開くことができる。ベンソン (Benson, 2014) は、「ナラティブ研究によって、しばしば個人の言語学習経験の視野を広げることができる。教室内外の環境や、様々な文脈的・心理学的変数を射程に収めるうえでも、他の手法より優れている。」(p.165) と述べている。このことから、ナラティブ的探究の利点は、研究者が他の質的データ収集方法の制約から解放され、参加者が経験を頼りに意味を構成していく方法を深く掘り下げることができる点にある。ベンソンのレビュー研究では、多彩な参加者からナラティブ・データを集めるための適切な道具がナラティブ・フレームであると説明している。このナラティブ・フレームを使って質的データを引き出す方法を初めて開発した人物はバークフイツェンとヴェッテ (Barkhuizen & Wette, 2008) であり、彼らは中国の大学で教鞭を執っている英語教員の経験を探究した。それ以降、多岐に及ぶ文脈で使用されてきている。例えば、ベトナムでは教育者の信念を調査したり (Barnard & Viet, 2010)、日本の大学生が外国で留学、労働した際の経験を測定したりするといった事例である (Kirchhoff, 2015)。ナラティブ・フレームの原則はひと続きの出だし文、接続語、文修飾語を与えて回答者を補助し、物語の限定的な特徴に集中してもらうことにある。標準的な質問紙とは異なり、学習者は予め決められた答えを選択するように縛られておらず、自らの経験を自らの言葉で自由に紡ぐことができる。合計で12の出だし文を作成し、様々な側面から学習のアセスメントを捕捉するとともに、複雑で多面的なデータを引き出せるようにした (表1参照)。学生には、現在受講している英語の授業で使用されたものの中から印象深いアセスメント手法を思い出してもらった。その際、総括的なアセスメント課題だけではなく、形成的な意味合いを込めた課題も含めることとした。

表1 ナラティブ・フレームで使用した出だし文^(注4)

-
- ・私の一番好きなアセスメントタスクは…でした。
 - ・なぜなら、…からです。
 - ・アセスメントタスクの中で、一番勉強になったのは…でした。
 - ・その理由は…からです。
 - ・一番難しかったアセスメントタスクは…でした。

- ・その理由は…です。
- ・この成績評価方法の良かった点は…です。
- ・この成績評価方法の良くなかった点は…です。
- ・学生はこのクラスでいい成績をとるためには、…ほうがいいです。
- ・なぜなら、…。
- ・もしこのクラスで成績が良かったら、…。
- ・しかし、もしこのクラスで成績が良くなかったら、…。

調査参加者

自己申告のアンケート調査では、日本の大学に所属する613名の大学1,2年生に任意で協力してもらい、552名の有効回答数があった。調査参加者の募集には、雪だるま方式の便宜的抽出法を採用した。合計で八つの教育機関が参加し、中部地方にある中堅私立大学の学生が主な参加者となった。回答者のうち、男性が279名、女性が272名であった。また、大学1年生が356名、大学2年生が195名であった。調査が行われたのは学生が授業の必修授業を受けている間であったが、農学、経済学、看護などの多様な専攻から構成されていた。学生には予め調査の目的と内容を説明し、匿名性を担保し、さらには、この回答が授業の評定に与える影響は何もないことを約束した。

ナラティブ・フレームのデータは、日本の3大学で開講された外国語としての英語(EFL)の八つの授業で収集した。合計で219名の学生がこのフレームに全回答した。回答にあたって学生には母国語で記入してもらうこととしたが、その理由は二つある。まず、質的データは極めて個人的なもので、そこから個人の深層の気質や態度を垣間みることができる。もし母国語を使用すれば、学生の思考から態度の表現へと結びつく軌跡を直接的に導出できるだろう。また、学生のほとんどが、必要とされる英語スキルを持っておらず、アセスメント経験を有意義で詳細に解釈することが難しかったためである。

データの収集と分析

調査結果の分析にはIBM社のSPSSバージョン22を活用した。記述統計量を算出した後、因子分析を行った。初期解を計算してデータ内の各因子の固有値を求めた。その結果、九つの因子の固有値がカイザーガットマン基準の最小値1を超え、それら分散の累積率は56.9%になった。しかし本研究では、ナラティブ・データも組み合わせながらデータを円滑に分析することを優先して、因子解を求めずに各カテゴリーに関連した記述統計量の値を用いるにとどめる。

もう一つのナラティブ・フレームは学期末の終わり頃、期末最終試験の1,2週間前に回答をしてもらった。重要なのは学生のアセスメント経験が鮮明な記憶として残っていることであり、夏季休業後に回答するのはあまりにも遅すぎる。この時期に実施したことで、学生は履修科目の最終成績に関する記述ができなかったものの、最終成績が彼らの感情に作用してしまうといった心配をすることなく、アセスメント経験を記入することができた。その後、回答を集計表に入力していき、日

本語で原文どおりに書き写した。このデータを研究者自身が入力した理由は、データに対する感度を高めて後の分析に役立てるためである。曖昧な言い回しや不鮮明な漢字(日本語)を見つけた場合は、日本語の母語話者に確認をしてもらい、明確にした。

ナラティブ・フレームの分析法には定性的内容分析を適用した(Schreier, 2012)。それは、ある程度の解釈を必要とする生データを、体系的に記述するのに適した方法である。この方法の核となるものがコーディング・フレームである。コーディング・フレームを構築する際の注意事項としては、次元性(メインとなる個々のカテゴリーは、生データの一つの側面だけを捉えなくてはならない)、相互排他性(切片化したデータは、一つのサブ・カテゴリーにのみ分類されなくてはならない)、網羅性(そのコーディング・フレームは、関連する全ての生データを説明するものでなくてはならない)、飽和性(各カテゴリーは、最低でも一度は使用されなくてはならない)、を担保する必要がある。そのうえで、ナラティブ・フレームの構造を演繹的ガイドとして参考にしながら、コーディング・フレームの次元を作成した。そして回答を分析し、帰納的にサブ・カテゴリーを導出していった。ここでは具体的に、マイリング(Mayring, P.)の包含手法を用いた(Schreier, 2012を参考)。次元を一度決定して、その側面に関連する事例を生データから探すように読み込んでいった。見つけた事例には仮のラベル名を与え、サブ・カテゴリーの印を付けた。事例の新たな情報が出てきたときは新しいサブ・カテゴリーを設け、サブ・カテゴリーで説明されている事例を見つけた際は、既出のコーディング・スキームで説明済みであることを補記した。このようにしてコーディング・フレームを開発した後、NVivo for Mac (v.11)に入力し、全てのナラティブ・データをコーディングしていった。

結果

量的・質的データから得られた結果は、文献レビューから判明したCHCの特徴に従って、以下のとおりにカテゴリー分けされた(注5)。

○ キャリアと社会進出

文献の報告によると、CHCの学習者は、アセスメントの良い結果とキャリアないし社会進出が深く関係していると思っている。本調査の結果もこの報告どおりであった。相対的に強く肯定された項目は「よい評価はよりよい職業につながる」(平均=4.23, 標準偏差=1.20)であり、それに比べるとやや弱い「評価が高ければ、わたしは社会でより高い地位につくことができる」(平均=3.84, 標準偏差=1.34)という項目も肯定的であった。

ナラティブ・データもまた、このアセスメント概念を支持している。19の回答で、授業で良い結果が得られれば優れた道具的価値になるだろうと、学生は記述していた。ある学生は、「もし授業で良い結果を得ることができれば将来良いことにつながりそうだから、その結果はとても役に立ちます」と述べている。別な学生はアセスメントの中でもビジネス会議のロール・プレイングを非常に気に入っており、将来の就職への備えになると考えていた。彼は、「それがテストだとしても、ビジネス会議で実際に会話をしているように感じました」と記述している。さらに8名の学生が、悪い

成績を取ると将来に響いてしまう、と書いていた。

○ 家族的恩義

文献による指摘では、CHCが特に重きを置くのは家族的恩義であり、成功の獲得は家族を喜ばせる一つの方法だといわれている。しかし、本調査データによると、アセスメント結果が自らの家族内の立場や、社会的にみた家族の地位にさほど影響を与えてはいないことが分かった。学生は、「学校でよい成績をとることが私にとって家族を喜ばせる一番の方法である」(平均=2.59, 標準偏差=1.44), 「家族は私が高い評価を得たときのみ、評価してくれる」(平均=2.80, 標準偏差=1.42)の項目に対して、否定的な見方を示していたからである。

この結果は学生のナラティブでも繰り返し確認された。回答には、「家族」「母親」「父親」、あるいは家族に関する概念が全くみられなかった。「もしこのクラスで成績が良かったら、」という出だし文を与えたところ、そこで表現された心情として最もよくみられた回答は肯定的な感情であり、「とても嬉しい気持ちになります」(回答数は65), 「もっと頑張ります」という意欲の向上がそれに続いた(回答数は33)。また、他の学生は「自信が高まります」(回答数は22), 「もっと楽しい気持ちになれます」(回答数は6)と記述していた。どのナラティブも、家族が喜ぶとか家族の地位が向上するといったことを示唆してはいなかった。同様に、「もしこのクラスで成績が良くなかったら、」という書き出し文を与えたところ、最もよくみられた回答は、次はいっそう頑張りたいという思いと否定的な感情だった。その中で、悪い成績を取ってしまえば多少なりとも家族の悪影響につながるだろう、と示唆した回答は一つもなかった。

○ 競争

三つ目の特徴として、CHCの教育システムにあると考えられているものは厳しい競争である。しかし、データを見るかぎり、アセスメント時の強いライバル意識はないと思われる。それどころか、学生はアセスメントの際に多かれ少なかれ仲間と助け合えると考えていた。「評価はクラス全体の協力を引出し、皆はお互いに助け合うようになる」(平均=4.20, 標準偏差=1.20), 「クラス内で評価されているときは、皆はいつもより助け合う」(平均=3.66, 標準偏差=1.13)といった項目について、学生は「そう思う」と答えていた。一方で、「そうは思わない」と答えていた項目には、「クラスで一番にならなければ、自分は価値がない」(平均=1.97, 標準偏差=1.20), 「評価が一番を選ぶための競争である」(平均=2.80, 標準偏差=1.35)があった。

次に、経験について書いたナラティブ・フレームを見てみると、学生は類似のコメントを記していた。例えば、32のコメントで、仲間と協働することでアセスメント課題が好きになったと述べていた。その中身は、「周りの人たちと話し合ったり、知らなかった言葉を互いに教え合ったりすることが、とても楽しかった」、「2人で一緒に頑張ることができた」などがみられた。複数の意見によれば、グループ学習のせいでアセスメント課題は難しくなったが、それはクラスメイトを前にした発表活動で不安が増したことと主に関係していた(コメント数は10)。大多数の学生はアセスメン

ト課題に対して友情と共助の精神をもって取り組んでいたようである。ある学生は、好きなアセスメント課題に挙げた理由について「友人と相談して答えることができたからです」と答え、別な学生は「知識レベルが他者と違っていても、何も心配せずに楽しく学べました」と述べていた。

○ 座学で得た知識か実践力か

上述したように、CHC は実践力よりも座学を大切にするという考えがよく知られている。本調査データはこの考えに疑問を投げかける。なぜなら、アセスメントが頭で分かっていること以上のものであることを学生は理解していたからである。「評価は本から学べることに限定されている」(平均 = 2.17, 標準偏差 = 1.11)の項目は、否定的な傾向が強かった。しかし、やや驚くべきことに、「評価は座学で得た知識のみに焦点を当てている」(平均 = 3.28, 標準偏差 = 1.23)に若干の肯定がうかがえた。どうやらここに、信念の葛藤が生じているようである。

学生が好んでいたアセスメント課題は、リアルな文脈で英語を活用でき、アセスメントが要求した様々な実践スキルの成長を確認できるものである。全体で39名の学生が、複数のアセスメント課題を授業で行うことに理解を示していた。「出席とテストとは別に、高得点を取るために別な方法があるのは良いと思います」という回答が一つあり、「(良いと思うのは)得点が1回や2回のテストで決まってしまうのではなく、いくつかの課題の組み合わせで算出されるということです」と言及していた。

量的データと質的データに相違があったことも特筆すべき点である。量的データに基づくと、アセスメントは信頼性があり、学習を追跡するうえで効果的だ、という点に強い同意がうかがえた。良い成績を取ることができれば、学習の努力が認められるとともに、自分が実際に紛れもなく成長したことを確認できるだろう、と学生は記述していた(回答数は19)。良い成績の原因は、アセスメントの仕組みがもたらした偶然だとか、教員が価値判断プロセスで何か失態をしたからである、という記述はなかった。唯一、学生の1人が「(高い成績をくれるのは)教員が私のことを気に入っているからです」と書いていた。同じく、成績が悪かった場合、それは取り組みが悪かったことを示唆しているのだと、学生は回答している。彼らの多くは、もし成績が悪ければ今後に向けてさらに努力をしようと思えるし(回答数は69)、その責任を自ら受け入れることもできる(回答数は17)、と詳述していた。ある学生は「私の英語運用能力は未熟なのだと思います」と、また、別な学生は「他の人よりも努力していなかったことを後悔します」とコメントしていた。

このことから、アセスメント結果が実際の能力と緊密に結びついていると学生は信じ込んでいる。それにまた、学習の正確な価値判断を妨害するアセスメント・スキームは、学生の不満項目の最上位にあった。得点を加算する方法の問題が不平不満の温床になっており(回答数は41)、得点の比重が偏ることで学習意欲が減退してしまう。そうした訴えは、「最終テストの配点が高すぎます」、「出席点の配点が低すぎます」などにみられた。他の学生は、アセスメントが不公平なせいで、作業を進めることが上手くできないと感じていた。それには、「友人がそれほどいない人には難しいと思います」といったものや、「テスト用紙を配布するとき、用紙を初めにもらった学生は早く取りかかれ

るのだから、不平等です」といった記述があった。したがって、学生は、アセスメントが基本的に信頼できるものだと認識している一方、公平性や平等性が考慮されていないことに神経を尖らせてもいた。

○ 実用性

CHC がもたらす五つ目の特徴とは、教育の本質が何よりも実用性にあるとみなす傾向である。本調査の回答はこの考えを支持しており、学生はアセスメントに有用性を求めていたことが示唆された。学生は「小テストは本番のテストの準備となる」(平均=4.46, 標準偏差=1.18)に肯定的であった。アセスメント結果を度外視することなく、「テストの結果は無視して捨ててしまう」(平均=2.48, 標準偏差=1.23)の項目には否定的であった。

アセスメント課題を有益だと考える理由として最もよく挙がっていたものが二つあった。それによって知識を成熟させることができるから(回答数は44)、そして、今のスキルを改善することができるから(回答数は40)である。学生によると、アセスメント課題が役に立つと考えるのは、例えば「正しい発音を学べる」とか「レポートの正しい書き方を理解できる」といったことに由来していた。また、語彙のアセスメントは「未知の語彙を使ってみることで語彙スキルが向上した」点で、口述のアセスメントは「英語文を自分で作成して説明することで、英会話スキルが伸びた」点で実用的であったという。また、19の回答で、授業の成績が良ければ、それが道具的価値になるということを暗示していた。もし高い成績を取ることができれば、十分な自信と能力を以って英検やTOEICのようなハイスイクスな総括的テストに臨めるだろう、と記述した学生もいた。さらに、多くの学生は自身のアセスメント・スキームを大切にしていた。なぜなら、「努力をすれば得点上がる」や「自らに課す努力は最終成績に反映される」のように、スキームによって努力が助長され、報われるものと思えるからである。

考察

CHCの主だった特徴には、教育における熾烈な競争(Carless, 2011)、家族集団への高い忠誠心(Hofstede & Bond, 1988)、実践スキルではなく座学で得た知識の優先(Han & Yang, 2001)がある。しかし日本においては、今回の量的・質的データがこの見解に疑問を呈している。まず、成績やアセスメントに強い競争心はみられなかった。調査参加者は、同じ授業の学生とよく協力し合いながらアセスメントができ、協働学習を楽しみ、学ぶ機会が増えたと報告していた。クラスメイトはライバルとして戦ったり負かしたりする存在ではなく、学習成果を高めるうえで益となる資源として、協力すべき存在だと理解していた。

また、アセスメント結果が家族関係に強く影響するという考えを学生は持ってはいなかった。アセスメント課題に懸命に取り組むうえで、家族を喜ばせる行為の優先順位は高くはなく、成績の良し悪しは家族内の評判とは関係がないと思われる。興味深いことにこの観点は、香港と中国の学習者が最も支持をしていない信念でもあることをブラウンとオウ(Brown & Wang, 2016)が突き止め

ている。試験に向けた努力に関しては、CHCの学習者が伝統的観念としての家族的恩義に従っていないという主張がますます支持される結果となった。

さらに、アセスメントは座学で学んだ内容だけで構成されるわけではないと、学生は考えていた。授業全体を通じて多種多様なアセスメント課題に取り組むことに理解が得られていたし、付言すれば、まだ出会ったことのない新規で斬新なアセスメント課題を学生は好んでいた。良い成績の獲得と、暗記・詰め込み学習やテスト文化に関連した学習方略に必要なスキルの習得との間にはあまり関連性がみられなかった。むしろ、能動的に授業に参加すること、様々な領域のスキルを向上させること、そして英語運用能力を実践的にリアルな状況の中で育むことが、良い成績を獲得することに結びつくのだと学生は解釈していた。本調査の参加者によれば、アセスメント課題に学習的な価値があると認められるかどうかは、その課題が実践的スキルと生きた知識を高めるものであるかが重要であった。

先行研究の主張によれば、日本の学習者にとって英語学習の真の目標とは、実践的なコミュニケーション能力であり、テスト受検ではない(Matsuda, 2011)。しかし、定期的な小テストやパフォーマンス課題は英語の実践力を高めるうえで良い方法であると、学生は理解していた。例を挙げると、ビジネス専攻の学生は、現実のビジネス場面で使いそうな言語を喜んで学習していた。また、農学を専攻する学生は、その分野に関連する語彙を学習することには価値があると思っていた。このように、現実世界で言語を使用することを想定した課題、さらに、学生がある程度選択できる課題が望まれていることは、すでに香港などの他のCHCでも確認されている(Carless, 2017)。これらを前提にすれば、今回の日本のデータは、儒教的伝承の背景を理由に学生はこのようであると決めてかかってしまうのは注意が必要であることを暗示している。

アセスメントの実用性という観点では、学生はCHCの観念と概ね一致するといえる。アセスメントが試験の準備に役立つ点には強い同意が得られており、アセスメントに対する総括的な志向性が示唆される。学生は、将来のキャリアや社会的地位と授業の成績が比例関係にあるとも捉えていた。同様の信念が香港の学習者にもみられており、彼らの主たる学習動機は、教育者や家族などの外部から認められることや、進学やキャリアの機会という意味で社会的に昇進することにあつた(Wang & Brown, 2014)。この観念は、学習者を努力主義に駆り立てるための効果的な外発的動機づけの装置といえるが、アセスメント結果をそのように認識してしまうと、真の成長という学習目標の価値を低下させてしまう危険がある。

フィードバックについては、アセスメント結果を無視したり無駄にしたりせず、得点に関心を持って何とかそれを活用しているようであった。アセスメントを形成的な目的で有意義に駆動させるための核心こそフィードバックであり、良いフィードバックによってより優れた学習成果に到達することはこれまでにも言及されてきた(Black & Wiliam, 1998; Ferris, 2010; Lopez-Pastor & Sicilia-Camacho, 2015)。しかしながら、世界で最も効果的なフィードバックがあつたとしても、それが使われずに無視されてしまうだけならば、誰のためにもならないだろう。すなわち、学習者はアセスメント結果に気を配ってはいるが、その結果と向き合い、解釈可能な状態にすることや、その結果

が教育者と学習者との健康的なコミュニケーションの上に成り立つものにはできるかどうかは、教育者の力量次第なのである (Hyland, 2000)。

日本のような CHC におけるピア・アセスメントの効果は、多くの諸研究が証拠づけているが (Taferner, 2008; Wakabayashi, 2008; Matsuno, 2009; Asaba & Marlowe, 2011; Sato, 2013), 本調査のデータによってその理由を部分的に説明することができる。学生は、アセスメントのおかげで授業が協力し合えるようになるという考えを支持していた一方で、過剰にライバル意識を燃やしたり過酷な競争を繰り広げたりして、成績で一番を取ろうとする考えを受け入れてはいなかった。現代の大学生が、学び合いやグループ学習を多く経験しながら成長していることが、その要因かもしれない。ゼロ・サム・ゲームとは呼べない教育システムにおいて、学生は協力やピア・サポートによる相互利益を意識できているのではないだろうか。

本研究の限界

本研究の分析データには限界がある。日本は儒教的伝承文化をもった国の一つでしかなく、他の全ての CHC でも同じ結論がみられると仮定することは早計であろう。考察の節では、CHC で実施されている諸研究の結果を可能なかぎり参照しているが、異なる文脈の諸研究も数多く利用して、CHC 諸国の学習者のアセスメント経験や概念の全体像をさらに鮮明に描き出す必要がある。さらに、データは幅広い専攻に在籍する学生 (例えば、農学、経済学、ビジネス) から収集されたとしても、英語の授業だけを対象としている。外国語の授業は、その本質からみても、教室の指導やアセスメントを含めて外国の文化的実践に取り組むことに近い。したがって、学習者の回答には外国語としての英語に対する感情が反映されており、異なる授業のアセスメント経験や概念ではまた違った回答になる可能性がある。

ペダゴジカルな展望

これまでの結果と考察から、CHC の学習者を相手に授業をしている教育者、あるいは CHC 諸国で働いている教育者に向けて、提言をいくつか行うことができる。以下に掲げる形成的アセスメント実践の提言は、英語圏の文脈の研究が支持している実践とよく合致している (Gardner, 2012; Sambell, McDowell & Montgomery, 2013; Jones & Saville, 2016)。カーレス (Carless, 2014) によれば、形成的アセスメント、すなわち学習志向のアセスメントは3本の柱を基礎にしているという。第1の柱は、学習志向のアセスメント課題である。日本と香港の CHC にいる学習者は、アセスメント課題とは事実上実践的なものであり、リアルな文脈での言語活用や個人の願いと密接に関連するものであると理解している。そして、そのようなアセスメントによって得られる実用性を大切にしている。この課題には、ロール・プレイングや、口頭発表、多読による作文やブック・レポートなどがあるだろう。学習者を強く動機づけているのは、社会的見返りを約束するアセスメントだが、それでも、学習における実質的な成長を向上させるアセスメントにも価値を置いている。学習者がアセスメント課題に取り組むときは、ピアでの作業や協働に極めて快く関与している。また、優れた

成績を得ようとするあまり、熾烈な競争やライバル意識を顕在化させることから生じうる負の影響は、微塵も感じられない。よって教育者に推奨すべきは、様々な種類のアセスメント課題を授業中に実施すること、特に、独創的で斬新な課題に関心を寄せることである。

第2の柱は、学習者の価値判断力を専門的に発達させることである。これは、競争意識の悪影響にあまり不安がることなく、学習者がピア・アセスメントや価値判断活動を実施することによって涵養されていく。学習者には、お互いの成果物を価値判断したり、授業の課題にコメントを残したりする活動を促すべきである。学生は、一緒に学習し、互いに学習支援ができることに楽しさや自信を実感していた。CHCの多くは個人主義ではなく集団志向的社会だとされるため、集団的学習への前向きな適応を後押しするかもしれない。グループのプロジェクト型学習や協力的な作業を伴う課題は、アセスメントを効果的にさせるだろう。

第3の柱は、学習者をフィードバックに関与させることである。学生の声によると、アセスメント結果は信頼できる指標として、真の学習到達度を示すものであった。その結果を無視せず、無駄にはしないと示唆されていたことから、豊かで詳細なフィードバックを提供することによって、その後の学習成果の向上に資することが期待できる。フィードバックの提供時にチェックリストやルーブリックを使用すれば、暗黙的になりがちな目標がより明確になるだろうし、今後の学習を方向づけることもできるだろう。また、日本と香港の学習者は、良い成績によって将来の社会経済的な見返りが得られると信じているため、効果的な外発的動機づけの装置としてテスト得点を活用し、より高い到達目標へ向かわせることができる。厳格な成績づけをすれば、学習者はより膨大な努力を投資することになるが、他方、過度に甘い成績づけをしてしまうと、おそらく学習者の努力は減退してしまうだろう。教育者の価値判断は、本物の能力を示す指標として、大切にされ、信頼されている。そのため、学習者がすぐに無視したり拒絶したりはしないことを教育者は知ったうえで、質的フィードバックを行う作業に時間と労力をかける価値はある。このようなテストのフィードバックと得点の使用は、カーレス (Carless, 2011) が総括的テストの形成的活用と呼ぶものである。それは、学習の促進を重視するアセスメントと、ハイスティクスな状況での順位づけと認定を重視するアセスメントの間に相乗効果を創る潜在力を秘めている。

結論

概して、本研究の結果が強調していたのは複数の潜在的問題であり、それは儒教的伝承文化のカテゴリーを使って、その国々にいる学習者がアセスメント課題にどのように取り組むのかを理解することで生じるものである。本研究において日本の学生が経験し、概念化していたアセスメントは、儒教的な教育システムでみられる主流な見方とは一線を画していた。それによれば、競争的なライバル意識がほとんどなく、アセスメント課題が実践力の成長を促すうえで重要だと理解しているのではないだろうか。学生の回答をみると、彼らは良い成績と努力が密接に関連していると信じ込んでいる。そして、良い成績を取ることができたのは、生きた実践力が向上し、ひたむきに应用させたからであり、他方、悪い成績を取ってしまったときには、全て自分のせいだとして、努力して次は

頑張ろうと考えるのである。とすれば、教育者は「CHCの学習者にアセスメント課題を与えても他の文化圏の学習者に比べて上手くいかないのではないかと心配する必要はあまりないだろう。いやむしろ、本研究と英語圏の成果が顕著なまでに類似していることを考慮すれば、これらの文化を比較したところで得られるものは少ないのではないだろうか。

(第2節については、第2著者が翻訳)

3. 解説, 今後の研究課題, 展望

当該論文は、日本の大学生が所有するアセスメント観の調査を通じて、既存の儒教的伝承文化の傾向性が、日本の学習者の気質とは必ずしも合致しないことを指摘したものである。補足として、資料1には実際に使用された6件法のアンケート調査項目を掲載している。

近年、形成的アセスメントが普遍的な技術ではなく、歴史性をもったペダゴジカルな文化的実践として解釈されるべきであるとする見方が強まる中、確かに、一連の混合法的研究は日本の大学生が抱えるアセスメント観を特定することに貢献し、こうした方法論に学ぶところは大きい。この社会文化理論に基づく分析視角は、より厳密に言えば、アセスメント活動それ自体に文化性を求めるか、それともその動機に求めるかに分岐しており、アセスメントの文脈的諸要因を探索する今回の研究は、後者に該当するだろう。一方で、論文内で言及された「本研究の限界」以外にも複数の疑義が散見される。以下、大雑把ではあるが、この論文における三つの課題を代表して取り上げたい。

一つ目は、学校段階や学習背景の未考慮である。アセスメントの経験や概念は、受験文化から解放された大学生と受験文化に苦しむ中学・高校生、さらに言えば、大学内でも就職活動をしている(控えている)学生とそうではない(就職に危機を感じない)学生では、異なってくるのではないだろうか。例えば、中学・高校生が塾通いのために親から莫大な投資を受けているならば、ライバル意識を高め、家族を喜ばせることが一つの動機になっていても不思議ではない。また、親がCHCの影響を受けた躰や家庭教育を施していた場合、(それを反面教師にできる場合もあるかもしれないが、)学習者も同じように導かれることが十分に考えられる。家庭での貧困の連鎖を断ち切るために「社会的な成功の証」を得ることを親から強く期待された学習者は、アセスメントに対してどれだけ肯定的な印象を持つことができるだろうか。すなわち、家庭教育や社会経済的背景などの諸要因が学習者のCHC的志向性を誘発しうるともいえるし、またその逆も仮定することができる。ある特定の社会的な価値や評判に肉づけされた大学に所属している1, 2年生の場合は、今回定義したCHCのアセスメント観が希薄でこそあれ、学習内容や専攻分野の違い以上に、こうした学校段階や学習背景を考慮せずにそれを一般理論として説明することはまだ距離があると思われる。

二つ目は、方法論上の問題である。そもそもC-SCoA自体、1因子あたりの項目数が少なすぎるものがあれば、項目同士の意味が区別しづらく、日本語に訳しても理解が難しい項目もある。また、回答者が協調やバランスの維持に根ざした《曖昧性》を思考様式として駆動させれば、質問項目の「だけ」や「のみ」といった限定用法的表現、あるいは「自分に価値がない」といった極端な断定に対して敏感に反応するはずである。このとき、彼らが《本音》と《建て前》を知っているとすれば、建て

前上の社会的望ましさを使って回答している可能性も排除できない。このように、質問の内容や尋ね方そのものが《文化的気質》を発動させる装置と化してしまっているのではないだろうか。さらに、当該研究は生データを既成カテゴリーに当てはめる方法を採用しており、新しいカテゴリーを創出するものではない。日本の文化的アイデンティティと C-SCoA で測定できる構成概念の一致度を理論的に検討してはならず、日本の文化的特色の可能性に関する議論が閉ざされてしまっている。

三つ目は、既存の形成的アセスメント論にある綻びである。確かに CHC は議論に開かれた概念ではあるが、形成的アセスメントもまた論争的な概念であり続けている。スタンダード達成のエビデンスとなるテスト得点の上昇に寄与するものとして、欧米諸国を中心にその効果が専ら検証されてきた現状に鑑みれば、「試験ではなくその代わりとなるアセスメント課題を重んじるとともに、教科書で学んだ知識ではなく実践的スキルの実質的な成長を促そうとする」(p.181) 役割が機能していることをどこまで断言できるだろうか。形成的アセスメントへの社会文化的アプローチが何を目的にしているのか、そして、提示した方法論がその目的を分析対象に含むものとなるのかについて明確な説明がなければ、「日本に直輸入される」英語圏由来のアセスメントの考え方や方法に内在する重大な問題が見過ごされてしまうことになりかねない。

以上3点の指摘より、日本と英語圏の文脈を比較することの意義に消極的な著者の論理には大きな飛躍があることは否めない。ここでは、日本にみられる多くの文化スクリプトと学習を支えるより広く深い周縁的要因が捨象されてしまっている。人間活動の歴史的発達の経過で出現し、社会の土壤に染み渡った CHC は、そこに生きる人間の思考を統御している実態を部分的にしか説明できないものである。今回の研究では、中国や香港と日本のアセスメント観の共通性と相違性が明らかにされたが、上述の課題を乗り越える作業に並行して、日本のアセスメント観をどのように説明できるのかが次なる挑戦といえるだろう。

とはいえ、以上の点は日本からの発信が少ないため、西欧圏からみれば無理からぬことかもしれない。そこで、日本の社会文化について少し深く考察する。特にこれは、学校教育に顕著であるが、CHC が顕著な大陸アジアと違って、CHC の他に、ハードワーク、掃除、祭りや課外活動、イベント、行事の影響(神道)、絆の社会的再生産(仏教)など、混和した状態(mixture)がある。さらに、誤解を恐れずに付言すれば、日本の大学教育実践には、大学院教育においても、「型」と「甘え」の文化があるとの指摘がある(McCarty & Hirata, 2010)。マッカーティと平田は、現在の高等教育改革を背景に、日本の教育哲学博士課程学生の伝統的な教育を、「自己教育としての学び」「他者との存在と学び」「職業への社会化」「外国語科目の学習」に焦点を当てて、西洋的・日本的な視点から分析している。特に、今井の「道具としての大人の自己」の日本的構築を説明していることは極めて興味深い。それによれば、日本の博士課程教育の重要な要素とは、集団に属し、社会階層の中で明確な位置を確立し、忠誠心、調和、尊敬の強い感情を育む親密な指導関係を築くという、職業上の行動様式(「型」)を学ぶことである。さらに、相互依存(「甘え」)を分析し、西洋の理想的な自律性と成熟性との対比も行っている。無論、コース特有の要因は抜きにできないが、こうした要因も加味しながら、より深い洞察を得ていくことができるのではないかと(注6)。

本学では、令和3年度からの新ISTU運用に伴い、ラーニング・アナリティクスが可能となり、形成的アセスメントとフィードバックが進む。それは、学生の思考活動と学習の変容・成果との複雑な関係性が解明される試みではあるが、その一方で、上記までに強調してきた学生目線の社会文化的特質への関心が、そこから抜け落ちてしまいかねない。その際に、日本の教育学ではまだあまり実績が見られないNVivo等の質的分析ソフトによるコーディング手法を取り入れ、エビデンスを学生から得ることも視野に入ってくると思われる。このことにより、具体の実践から一般理論を生成していこうとする運動を活性化させ、実践と理論を架橋する方法論的枠組みに関心が寄せられるようになれば、教育学研究方法論自体のイノベーションを図ることができる。結果として、大学院生の研究方法論の深化と業績の向上、博士特別研究員等の増加・質の向上をエビデンスとして示すことを目指すこともできるだろう。さらには、このエビデンスを携えて、アジアの大学でのエンゲージド・ラーニング向上の議論に与ることが期待できる。

最後に、エンゲージド・ラーニング向上の議論と関連して、課題克服と展望を実現させる一つの足がかりは、いくつか考えられる。その一つとして、第2言語の習得という科目にこだわらず、もともとは教員側が経験的に振り返るための変数の集合体であったが、アクティブ・ラーニングに関連した授業評価項目(案)を資料2に掲載した。これを、学生に回答してもらうことで、教員が授業改善を目指す道具となるものであろう。項目を精緻に検証し、議論の練り上げが強く求められよう。

(第3節については、第1著者と第2著者が共同執筆)

謝辞

本研究資料の作成にあたり、論文と資料のご提供を快諾してくださった Paul Wicking 先生(名城大学外国語学部准教授)に感謝申し上げます。なお、本研究資料の作成、執筆にあたっては、本研究科の2020年度国際共同推進事業経費の助成を受けた。

【注】

1. 「社会的背景」の節名は訳者が付記した。
2. 下線は著者による(以下、同じ)。
3. C-SCoA は全33項目あり、当該調査は項目「アセスメントは、社会的地位に大きな影響を与えるものである」を欠いている。
4. ナラティブ・フレームの調査用紙は原文のとおりである。
5. 以下、斜字体と質問項目の日本語訳は、調査用紙に書かれた原文のとおりである。ただし、タイトルでアセスメントと評価(エヴァリュエーション)とあるように、グレーディングなどの成績評価とは異なる、価値判断を伴わない認識、観察、解釈というアセスメントのフェーズを区別することは、重要である。
6. より分析的な見解では、アセスメントや評価に与えている学習環境や倫理観、文化の影響力を考慮できるよう、エートス指標の質問紙の改訂版を提案したバーデンの研究がある(Burden, 2008)。

【参考文献】

- Burden, P. (2008) . The use of 'Ethos indicators' in tertiary education in Japan. *Assessment & Evaluation in Higher Education*, 33(3): 315-327.
- McCarty, L. P. & Hirata, Y. (2010) . East Meets West in Japanese Doctoral Education: Form, Dependence, and the Strange. *Ethics and Education*, 5(1): 2741.

資料1

Student Conceptions of Assessment Survey

質問1～32は、もっとも当てはまる数字を一つ選び、○をつけてください。(6件法)

1. 家族は私が高い評価を得たときのみ、評価してくれる
2. 評価対象となる活動をするときは、クラス全体がお互いにより協力しあう
3. 評価は信頼できるものである
4. わたしがどのような評価を得るかは家族にとってたいへん重要なことである
5. 評価は、仕事や教育のためにもっとも優れた人を選ぶために行われるものだ
6. 評価が高ければ、わたしは社会でより高い地位につくことができる
7. よい評価はよりよい職業につながる
8. 私の成績は家族に対する世間の評価に影響する
9. 評価は一番を選ぶための競争である
10. 評価は決して終わりが無い。それは生きることの一部だ
11. 学校でよい成績をとることが私にとって家族を喜ばせる一番の方法である
12. 成績評価は十分に正確である
13. 資格試験の結果でその学校のレベルがわかる
14. 成績評価は学業以外の自分の長所を考慮していない
15. 評価に関係する活動をクラス内で行うときには、クラスの雰囲気がよくなる
16. 社会の中で成功するかは、どのくらい高い評価を得るのかにかかっている
17. 試験で満点を取るか、一番にならないと自分は賢いとは言えない
18. 評価は座学で得た知識のみに焦点を当てている
19. 教師は次に教えることを知るためにテストの結果を使う
20. 小テストは本番のテストの準備となる
21. 資格試験の結果は学校の価値を決める
22. 社会での高い地位は学校の成績と直接関係している
23. クラス内で評価されているときは、皆はいつもより助け合う
24. クラスで一番にならなければ、自分は価値がない
25. クラスの他の人や友達自分より高い評価を得ている
26. テストの結果は無視して捨ててしまう
27. テストの結果は保存するが、無視する
28. 学校での評価は社会や家族に対する自分の価値を決める
29. 大人になるということは、定期的に何らかの評価を受けるということである
30. 評価はクラス全体の協力を引出し、皆はお互いに助け合うようになる
31. 評価は本から学べることに限定されている
32. 教師はわたしの進歩を評価で確認することができる

資料2

アクティブ・ラーニングと授業改善に有益な評価項目(案)

次のうち当てはまる回答を丸で囲ってください。ただし、2から8については複数選択可。

1. 授業の中で、積極的に学ぶことを目的としたアクティブ・ラーニング活動はありますか？

あり, なし, 知らない

2. アクティブ・ラーニングなどの能動的な学びを育むために、教師は授業の中でどのような活動を行ってきましたか。授業内と授業外の活動の両方を考慮してください。

ブレインストーミング, ライティング, クイズ, インタラクティブな講義,
大人数でのディスカッション, ケーススタディ, ロール・プレイング, フォーラムシアター,
グループ・プレゼンテーション, ピアレビュー, 振り返りの時間,
体験学習(現場見学), その他

3. どの方法がより良い学習経験を与えることができますか？

ブレインストーミング, ライティング, クイズ, インタラクティブな講義,
大人数でのディスカッション, ケーススタディ, ロール・プレイング, フォーラムシアター,
グループ・プレゼンテーション, ピアレビュー, 振り返りの時間,
体験学習(現場見学), その他

4. 問題解決のために、自分を主体的にさせる効果的な方法は何だと思いますか？

質疑応答, クイズ, ケーススタディ, 大人数でのディスカッション, その他

5. アクティブ・ラーニングに適した雰囲気づくりの方法論は何だと思いますか？

問題解決型学習, プロジェクトベースの学習, 事例ベースの学習,
ブレンドラーニング(対面とオンラインの併用), フリップした教室(反転学習), その他

6. 授業でのアクティブ・ラーニングの取り組みについて、どのように考えていますか。

非常に満足している, 満足, 満足も不満もない, 不満, 非常に不満, 知らない

7. アクティブ・ラーニングの取り組みの効果をアセスメントするためにはどれが最適だと思いますか？

学習成果の測定(試験, テスト, 課題の成績など), 学生アンケート, 学生インタビュー,
生徒のフォーカスグループ, 観察, 課題・活動の完了数, その他

8. アクティブ・ラーニング活動の前に必要と思われる情報はどれだと思いますか？

教員が活動を説明する(目的, 文脈, 根拠, どのように行うべきかなど),
教員が受講生にコースの構成(シラバスの内容)を伝える,
学生が, この活動に対して何を期待するかを説明する, その他

9. 知識を学ぶためだけであれば、どちらが良いですか？

学習者中心の授業, 教授中心の授業, 双方ともに知らない

10. 技能を身に着けるためには、どちらが良いですか？

学習者中心の授業, 教授中心の授業, 双方ともに知らない

Contributing to Educational Assessment and Evaluation in the Field of Higher Education through Data Collection

Masahiro ARIMOTO

(Professor, Graduate School of Education, Tohoku University)

Kohei NISHIZUKA

(Graduate Student, Graduate School of Education, Tohoku University/JSPS Research Fellow)

Classroom assessment in our department has been carried out for more than 10 years, but it is done routinely, based on the idea that faculty members have to do the teaching evaluation, and there is no concept of enjoying the students and the assessments themselves, so it has not produced any significant results. This is an area where there is a big gap between our school and other countries. With our graduate school now online, it has become much easier to obtain and organize and analyze data, and due to the change in the balance of power, the demand from students for innovation in lectures and learning has become stronger, and the awareness of faculty members has reached a certain point of change. However, there has been almost no dissemination in English in the field, especially from domestic schools of education, graduate schools of education and other universities. In this context, with the kindness of a third author, who has demonstrated the presence of assessment through data collection using mixed research methods in the peer reviewed international journal , we will examine the supplement and commentary of Wicking (2020) . We propose this paper as a cornerstone for further international collaboration, listing the challenges. It would also be applicable to Faculty Development (FD) .

Keywords : Formative assessment, Confucian heritage culture, learning-oriented assessment, Japan, English as foreign language