

キャリア教育科目における学修評価の課題

ーパフォーマンス評価とルーブリックの活用可能性を求めてー

Evaluating University Career Oriented Education: Using Rubrics for Performance Evaluation

和田 佳子

WADA Yoshiko

In recent years, assuring high-quality education has become essential for universities. In addition, the status of university career oriented education courses has gained importance with regard to connecting new graduates to society as well as the professional world. Universities must enrich the content of career oriented education, but they require effective course evaluation methods to do so.

This paper investigated the status of career oriented education at universities in Hokkaido. Most career education classes employ an active learning style, which has been demonstrated to provide fair results; however, it makes rigorous performance evaluations difficult. Thus, we need to explore performance evaluation systems for university career education courses, which lack reliability and validity. One method to overcome these problems is using a rubric, a recently introduced educational assessment tool. Using a rubric not only enables rigorous evaluations but also presents advantages to those who are evaluated.

In a case study of an evaluation of volunteer activities, I discussed the nature of assessment and the desired validity of the evaluation. However, further scientific study of evaluation methods is required to improve evaluators' awareness and ability and provide students the advantages of excellent evaluation in career oriented education (and other) courses.

はじめに

日本の大学を取り巻く環境の変容は著しい。グローバル化に対応し得

る高等教育機関の「質保証」が重要な政策課題となり、日本の大学は世界に遅れをとりながらも、社会の要請に応え得る大学の本質的な在り方を模索している。特に2008年の中央教育審議会答申「学士課程教育の構築に向けて」では、それぞれの大学が独自の教育理念と到達目標を学内外に明示するとともに、学生の成長を実現させる場としての学士課程教育の体系化と成果の「質保証」を厳しく求めている。「学士力」として在学中に学生が身につけるべき具体的な能力・技能は何なのか、成長や学修成果をどのように評価・担保して社会に送り出すのか等、今、日本の大学には、社会が求める教育水準の「質保証」を最低限の責務として突きつけられている。

米国では、2006年9月に教育省のマーガレット・スペリングス長官が、米国の大学のラーニング・アウトカムズに関する基準に見合った教育を行っているかどうかを各大学に数値で示すことを要求し（スペリングス・レポート）、その余波は日本にも及んで、教育の質保証および評価の厳格化・可視化への動きにつながった。また同時期に、経済産業省の「社会人基礎力」、文部科学省の「就業力」、厚生労働省の「就職基礎能力」等の、社会人育成指標およびコンピテンシーが具体的な形で次々と公示されたことは、育成基準や到達目標の共有化を促すひとつの契機になった。

最近では、新卒採用時に大学を限定し、社会人基礎力指標とコンピテンシーを採用条件に入れるという企業も少なからず見受けられるが、これは採用リスクを抑えると同時に大学に質保証を求める動きのひとつであるとも考えられる。さらには卒業時点の能力保証にとどまらず、教育の成果をより長いスパンで保証することも求められている。文部科学省は既に卒業生を追跡し、中長期の教育効果の検証を行う、「学校から社会・職業への移行に関わる縦断調査プロジェクト」に着手しており、大学の質保証の名の下に、ラーニングアウトカムに対する責務は増大する一方で

ある。

以上のような背景から、基礎力を確実に固めるための初年次教育と、出口の水準を担保するためのキャリア教育が、これまでの大学の一般教育や専門教育（学部教育・アカデミック教育）を補完するものとして強化されることとなった。2011年の中央教審議会答申で、教育課程の内外を通じた「社会的・職業的自立に向けた指導（キャリアガイダンス）」が法令化（平成23年4月1日施行）されたことによって、日本の全ての大学が、学校社会と職業社会の接続に重点を置く教育の実現を目指す動きにある。

本稿では、大学教育の質保証が問われる時代の、教育評価の動向を捉えた上で、キャリア教育評価の特徴であるパフォーマンス評価および真正評価の課題を探る。その際、本学学生が参加したボランティア実習の評価を事例にして、評価の妥当性と望ましい評価の在り方について考察する。

1. 高等教育機関の出口の質保証とキャリア教育

大学における出口の質保証ニーズの高まりによって、一般教育科目・専門科目と架橋し、社会への接続の調整機能を果たすキャリア教育が、大学教育全体の成果保証の一翼を担うというポジションがようやく定まりつつある。しかしながら、現実には学士課程内での体系化は未整備などころが多く、教育内容の深化や教育技法の改善、成果の測定と検証などそれぞれの精緻化はこれからの大きな課題である。

キャリア教育とは、一人ひとりのキャリア発達を支援し、それぞれに相応しいキャリアを形成していくために必要な意欲・態度や能力を育てる教育、また学生が卒業後自らの資質を向上させ、社会的および職業的自立を図るために必要な能力社会的自立を促す教育⁽¹⁾のことである。前

述のとおり 2011 年以降, 高等教育機関では教育課程内のキャリア科目設置が推進されてきた。また同時に, キャリアセンター等が主催する正課科目外のキャリアガイダンスについても内容・手法ともに高度化が進んでいる。

筆者らが行った道内私学文系 10 大学のキャリア教育の現状ヒアリング調査 (椿・和田 2012) では, 対象 10 大学全てにおいて, ここ 3~4 年以内に全学生が履修できる位置にキャリア科目を開講し単位認定を始めていることが確認された。また, 有資格の就職課専門職員やキャリアカウンセラーの任用により, キャリアガイダンス・プログラムの内容は格段に充実し, 教授法の高度化も確認され, 学生に対する就職およびキャリア支援の体制は遍く整ってきたことが窺われた。

大学における出口の質保証について, これまで日本の産業界では, 即戦力となる人材は産業界で育成する, つまり企業で必要とされる人材育成は入社後に企業が担うという暗黙の了解があり, 大学教育の出口成果が厳しく問われることは稀であった。極論になるが, 大学から社会への移行接続において, どのような質の学生を送り出そうとも, 表立って大学の責任が問われることはなかった。しかし現在では高等教育機関に対して, 社会で通用する一定水準の知識やスキルを備えた人材を在学中に育成し輩出すべきとの社会的圧力が高まっている。しかも, それは卒業時点の就職内定率の高さを意味するようなものではなく, 卒業後も中・長期にわたって継続的に効力を発揮するカレッジ・インパクト⁽²⁾ でなければならないのである。

当然のことながら大学卒業時のキャリア教育目標の達成度 (平たく言えば, 仕上がり具合) は, キャリア教育科目単体で実現するものではなく, 大学内の全ての授業, 課外活動, さらには学外におけるアルバイト, インターンシップ, 地域活動, ボランティア活動, 留学など様々な体験と人々との関わりに影響されて醸成されるものである。社会の扉の前に立

つ時にあるべき像を想定して、大学生生活全体を統合するのがキャリア教育の役割であると考ええる。

2. 高等教育機関におけるキャリア教育の評価の現状と課題

近年の大学のキャリア教育科目の実施状況（和田・椿「道内大学のキャリア教育とインターンシップの現状調査」2012）結果では、より実社会に近づけた実践的教育を施す必要性から、座学よりもアクティブ・ラーニングによる教育技法を用いた科目が目立ち、今後も PBL（Project-Based Learning）など課題解決型授業が増加することが予想される。学生参加型のグループワーク、地域体験学習、インターンシップ、ボランティア活動、フィールドワークなどのアクティブ・ラーニングは、嫌が応にも学生が主体的に行動する状況を生み出し、自らが動くことで失敗体験をも含めた幅広い学びを獲得し、結果的に達成感と成長実感を得やすいものである。運営の仕方次第ではあるが、座学による受動的な学びと比べて、大学で主体的に学ぶことの面白さや重要性に気づききっかけになる可能性が高いため、継続的な教育効果の高まりにも期待ができる。

しかし、これらの学修成果をどのように測定評価することが望ましいのだろうか。キャリア教育ではリアリティある社会的体験などを特徴とするため、パフォーマンス評価や真正評価が主流となる。それ故に評価は単純ではない。例えば、出席状況が良好で定期試験の点数が高い学生が職場でそのまま通用する能力を備えているとは限らない。その場合、何をどのように評価することが妥当と言えるのだろうか。評価の対象、実施時期、評価の妥当性など課題が多々存在しているが、これらについての先行研究は未だ稀少である。

正解がひとつではない科目特性から、基軸を曖昧にしたまま評価者の裁量によって、経験的な尺度や勘を頼りに評価している例もないとは言

えない。評価する側・される側双方が納得する評価とはどういうものなのか、厳密性と質保証に応え得る評価の在り方を検討していくことが必要である。

大学における成績評価の根拠となる大学設置基準第25条の2では、「評価の客観性及び厳格性の担保と、基準の明確化」が謳われている。また、評価の厳格化の必要性は平成22年の学校教育法施行規則の改正（23年4月施行）でも、「学生に対してその基準をあらかじめ明示し、当該基準にしたがって適切に行う」として念押されている。

成績評価の厳格化が声高に叫ばれるようになった背景には、大学の成績評価の信頼性の低さという問題がある。2008年の中教審答申では成績評価について、「個々の教員の裁量に依存しており、学生確保という経営上の要請とも相まって、なし崩し的に安易な成績評価が広がるおそれがある」ことが指摘された。また、「教員間の共通理解の下、各授業科目の到達目標や成績評価基準を明確にするとともに、GPAをはじめとする客観的な評価システムを導入し、組織的に学習の評価にあたっていくこと」に触れている。

キャリア教育をテーマにして、昨年、東京で開催されたあるシンポジウムの席上で、パネラーの一人である大手企業採用担当者が、職業社会の基準で見るとGPAを導入しようがしまいが、大学教員による成績評価ほどあてにならないものはない」と明言し、続いて別のパネラーが、「もともと大学の成績は見えていない。企業内で独自の指標を作って判断している」と発言し、会場内にいた大学関係者に動揺が走った。つまり、大学の教員がつける成績の高さと仕事能力の高さは相関しないということである。教育界と実業界の立場の隔たりの大きさを認識しつつも、少なくとも、大学におけるキャリア教育（科目）の評価と仕事で活かせる能力のレベルを接近させること、そのため評価手法について研究を進めることは当然

ながら、「質保証」の責任上、まずは評価者の意識レベルを高めることが先決である。

学修評価には大きく分けて直接評価と間接評価の2つの型がある。間接評価は学生が、学生調査やアンケート、インタビュー、レポート等を利用してどのように自己認識して行動につなげたか、学修の成果を間接的に評価することである。いわば学生自身が「何ができると思っているか」を測るものである。一方、直接評価とは、正誤が明快な客観テストにより学生の知識や行為を通じて成果を直接的に評価し、「何ができるか」を測ることである。

さらに、この直接評価の中には、学習者自身の作品や実績、発表などを直接に評価するパフォーマンス評価と、学修者の成長を測るためのポートフォリオ評価、職場や市民生活など本物らしさをもった課題に取り組みさせて評価する真正の評価などがある。

本研究のために収集したキャリア教育の評価に関する先行研究から、キャリア教育の評価方法をブルーム（Bloom.B.S）の診断的評価（diagnostic evaluation）、形成的評価（formative evaluation）、総括的評価（summative evaluation）に当てはめて、以下のように分類してみた。

（1） 職業観・就業観を基準にした評価

キャリア意識、キャリア形成態度についての省察。レポートの記述語からのテキストマイニング、面接評価など入学から卒業までのデータによる評価（寺島 2009）〈診断的評価〉

（2） 能力・コンピテンシーを基準にした評価

「社会人基礎力」等による各項目の達成度チェック、コンピテンシー・チェックリストの活用による評価（内浦・毛受 2008）。〈診断的評価〉

（3） 自己効力感を基準にした評価

セルフ・エフィカシー（Bandura 1977）をベースにした「進路選択に

対する自己効力尺度」(浦上 1996) などの活用 <診断的評価>

(4) 学習意欲を基準にした評価

学生自身による, 取組み態度・意欲についての自己評価。客観的指標としての受講状況データなど。<形成的評価>

(5) 進路選択・就職との関連にもとづく評価 (寺島 2008)

在学中のラーニングポートフォリオ, 進路選択時の各種効果測定結果, 行動変容など, 入学時から卒業時までのデータ収集による評価。<総括的評価>

さらに, 今後の動向を見通して付け加えるとすれば, 生涯にわたるキャリア形成に影響を及ぼす「中・長期的な評価」(卒業生追跡調査) <総括的評価>が追加されようか。

3. キャリア教育におけるパフォーマンス評価と真正評価

前章で述べたとおり, キャリア教育科目の評価にあたっては, 知識の量や正誤を測り評価を付けるだけでなく, 学生ひとりひとりの「パフォーマンス」に対して評価を付けるウェイトが大きい。特に近年は, 課題解決型学習など学修者の自律性を重視する授業が実施されるようになってきており, この傾向はますます高まると思われる。例えば PBL に代表されるようなグループ学習では協働のプロセスを重視するため, 最終的な成果の評価だけではなく, 過程も評価の対象となる。その場合, 公平性や妥当性に配慮しながら真正評価として多角的に判断する必要があり, 評価者・学修者双方が納得できる評価を付けることは容易ではない。

筆者が所属しているビジネス実務学会では, 早くから学生のプレゼンテーション教育技法の研究に取り組んでおり, その評価方法についての議論を重ねてきた。またインターンシップやプロジェクト型学習, ビジネスマナーの習得度についての測定方法など, いわゆる学生のパフォーマ

ンスに対してどのように評価をつけるべきかの議論を繰り返してきたが、科学的論証に到達しているものはまだごく僅かである。

このようなキャリア教育における学修者の活動、評価の対象（何を評価するのか）、評価基準や評価の時期などの課題をクリアするために、まずはその基盤になるパフォーマンス評価の原則を押さえておく必要がある。

パフォーマンス評価とは、ペーパー上に書き込まれた解答を評価するのではなく、学生の発表やレポート、資料や作品から思考力や判断力、表現力などを判定評価する評価法のことである。教育学の中ではパフォーマンス評価についての研究の歴史は長く、とくに初等・中等教育機関では既に様々な試みがなされている。パフォーマンス評価の方法にはアンケート調査やインタビュー分析、チェックリストによる自己評価、記述式による振り返りシート記入、相互評価（ピア評価）などがある。大学生の場合は特に、学修者の自律性を促す目的で自己評価を加味することが望ましく、その際、教員の評価が学修者本人の評価と大きく食い違うことがないように留意しなければならない。高等教育に先んじて取組みが始まった小・中等教育機関のキャリア教育のパフォーマンス評価研究の蓄積を援用し、高等教育機関に適した評価方法の構築が急がれる。

4. パフォーマンス評価のためのルーブリックとは

パフォーマンス評価の精度を上げるために活用が考えられるもののひとつに 1980 年代に米国で開発された評価ツール、ルーブリックがある。ルーブリックとは、評価規準を明確にし、学生が到達レベルを確認することを目的に使用される評価ツールである。米国では AAG (米国大学協会) が VALUE RUBRIC として作成し、米国の大学の質保証の一環として広く活用されている。

日本では初等・中等教育機関で、ウィスコンシン州、アルバーノ大学のルーブリックなど代表的なモデルを参照しながら、研究導入された経緯が見られる。松下は小・中等教育の評価研究として早くからルーブリックに着目し、フィギュアスケートや体操のパフォーマンス評価に例えてルーブリックをわかりやすく説明し、日本の大学での普及を勧めている。

ルーブリック評価は、評価者と被評価者の双方にあらかじめ評価規準と評価基準を提示して評価の観点を可視化するものである(表1)。日本の大学では、関西国際大学などで先進的に導入が始まっている⁽³⁾。中教審大学教育部会委員でもある関西国際大学の濱名はルーブリックの効用を詳しく説明している(2011)。濱名の説明資料によれば、ルーブリックを用いることで公平性、平等性を担保することにつながり、学生の側からすれば「学習者として、どのように評価され、成績を付けられるかがわかる」、「評価によって学生自身の成長を確認できる」、「自らの学修活動の強み弱みの理解の助けになる」という。濱名によるルーブリックの説明資料から特徴と効用を抜粋すると、以下のようになる。

- ① ルーブリックとは学生が何を学習するのかを知り、評価基準と学習到達レベルを示すものであり、具体的な評価基準をマトリクス形式で示す評価指標である。
- ② 学習者の達成度合いを示す数値的な尺度 (scale) と、それぞれの尺度に見られるパフォーマンスの特徴を説明する記述語 (descriptor) で構成される、評価基準の記述形式として定義される。
- ③ 記述語とは、評価の視点、あるいは観点にあたるもので、尺度というのはレベル、段階的水準といえるもの(高浦勝義 2004『絶対評価とルーブリックの理論と実際』より)のことである。

- ④ 「評価軸」を示しておことは、「何が評価される事柄なのか」についての情報を共有するねらいもある（熊本大学ウェブ「学習指導・評価論」より）。
- ⑤ OECD/PISA の国際調査「生徒の学習到達度調査」や文部科学省の教育課程実施状況調査でも一部、ルーブリックは導入されている。

上記のような特徴を持つルーブリックは大学の授業の評価、特に正誤だけを問う性質の科目ではないキャリア教育評価に有効であると考えられる。ルーブリックの例（表2）のように、具体的な記述語に沿って評価することで、評価者の迷いを少なくする効用があり、結果のフィードバックによって、学生の到達度合いと不足度合いを客観的な情報として具体的に伝えることが容易になる。また、ルーブリックはグループ学習の場合にも、結果を他者と共有してメンバー間の調整に役立てることもできる。さらに、「何を目指して、今何をやっているか」が、時として曖昧になりがちな課題解決型学習においても目標や内容を常に確認できるなど、ルーブリック活用のメリットはある。

しかし一方で、ルーブリックの作成には手間と時間がかかるという欠点がある。完璧な記述語を載せたマトリクスを作成することに没頭し、「評価をつける」ことが目的になってしまうことがないように気をつけなければならない。ルーブリックを用いて行う評価行為は、学修者の成長を促すためにあるという、元々の目的を忘れることがあってはならない。

5. ボランティア実習におけるパフォーマンス評価の事例検討

ここではキャリア教育におけるパフォーマンス評価についての考察材料のひとつとして、平成24年度に本学で1年次に開講された「ボラン

表1 ルーブリック作成の基本形式

	S	A	B	C
評価項目1	評価基準	評価基準	評価基準	評価基準
評価項目2	評価基準	評価基準	評価基準	評価基準
評価項目3	評価基準	評価基準	評価基準	評価基準
評価項目4	評価基準	評価基準	評価基準	評価基準

S:super (期待する以上のプラスアルファが見られる) A:十分に目標を達成しているレベルである B:おおむね満足できるレベルであるが、未到達な部分もある。C:努力を要するレベル (期待する思考活動が見られない。目標を達成したとは言えない)

〈参照〉 黒上晴夫、見える「評価」で授業が変わる～ルーブリックで授業作り
http://www.justsystems.com/jp/school/academy/hint/rubric/ru01_01.html (2013.1.10 取得)

表2 ルーブリックの例：(ボランティア実習評価を例に筆者作成)

評価の観点	S	A	B	C
取組みの姿勢	期待以上に主体的に取り組み、現場の求めるレベルに十分達している。	主体的に取り組み、現場の求めるレベルに達している。	おおむね積極的に取り組んでいるが、場面によって物足りなさがある。	積極性に欠け、現場で通用するレベルには達していない。
	追加コメント	追加コメント	追加コメント	追加コメント
課題発見・解決力	期待以上に自ら進んで仕事を見つけ解決する場面が多々あった。	自ら進んで仕事を見つけ解決する場面があった。	ある程度、問題に気づくことはできたが、自ら解決するには至っていない。	問題に気づくことができず、自ら課題を解決する場面はなかった。
	追加コメント	追加コメント	追加コメント	追加コメント
周囲との協調性	常に周囲の人と積極的に関わり、バランス良く行動し、職場における協調性は十分である。	周囲の人と進んで関わる姿が見られ、協調性が感じられた。	周囲の人とう関わろうとする努力は見たが、取り立てて素晴らしいと言えるレベルではない。	積極的に人と関わろうとせず、孤立する場面が見られた。
	追加コメント	追加コメント	追加コメント	追加コメント

ティア II (実習)」を例に取り、実習終了後に行った学生調査をもとにその評価課題について考察する⁽⁴⁾。

平成 24 年 7 月～8 月、夏季休暇中の 4 日間を利用して、社会学部地域社会学科学生 44 名が、16 か所の事業所に分かれて、ボランティア実習に参加した。事前指導として、「ボランティア I」(前期 2 単位)で、地域社会や福祉の課題についてのレクチャーを受け、規律やマナーにまつわる講話と実技の講習を受けて実習先に向かった。

実習終了後は、「社会人基礎 V」の講義時間を利用して体験の振り返りを行い、さらには情報共有と成果報告を目的とした実習体験プレゼンテーション・コンテストを実施して収束とした。これら一連の活動の最後に行った学生調査(アンケート調査)をもとに、実習後の学生の自己評価と実習先から得た外部評価の関係を検証し、パフォーマンス評価の妥当性について考察した。

【調査の概要】

- (1) 調査日時：平成 24 年 10 月 30 日「社会人基礎 V」授業終了時
- (2) 調査対象：実習参加者 44 名
- (3) 回収率：91.0% (40 名)
- (4) 調査内容：①実習参加前の自己目標、②実習先希望度、③実習中の行動についての自己評価、④受け入れ先評価の妥当性、⑤報告会準備・発表時の自分の役割と貢献度など
- (5) 結果の考察：
 - 1) 実習前の目標設定

本調査によれば、実習前に学生各自が掲げた目標は以下のとおり(図 1)、主に「意欲的行動維持目標」(意欲的に行動する・粘り強く行動する、やる気を維持する)と「対人積極性維持目標」(対人関係で悩まない、ネガティブ感情を出さない)の 2 つに集約され、他に「自己分析推進目標」

(強み・弱みの発見, 自信の強化), 「ストレス耐性強化目標」(落ち込まない, 細かいことで悩まない) と続く。この回答者集団を対象にして別の時期(1年前期・6月)に行われたEQ検査(就業能力適性に関わるアセスメントのひとつ)の結果, 「対人積極性」と「行動積極性」の弱さを特徴とする集団であるというデータが学生たちにフィードバックされていたことに照らし合わせると, 本人たちが自分の弱点を自覚的にとらえて目標設定し, 実習に臨んでいた可能性がある。

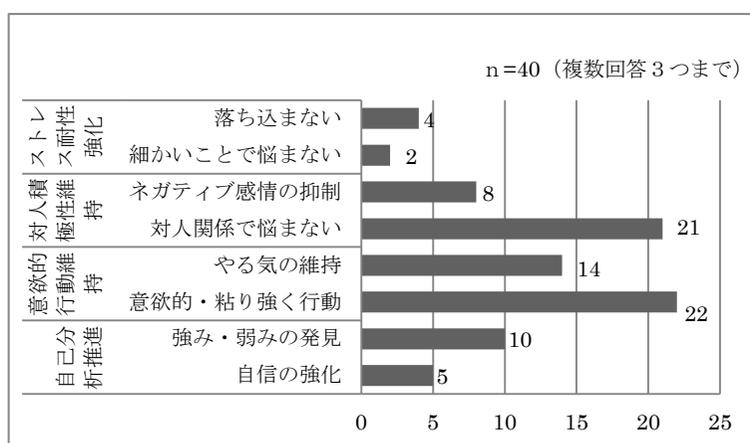


図1 ボランティア実習前に立てた目標

2) 実習先に対する希望度合い

事前に大学で準備されていた16か所の受け入れ先については, 原則的には学生の希望を踏まえながら振り分けが行われたが, 必ずしも第一希望であったとは限らない。実習中の行動との関わりを見る上で, 希望度合いを尋ねた。その結果, 8割が「希望に叶っていた」と答え, 「全く希望に叶っていない」との回答は7.5%(3名)であった(図2)。

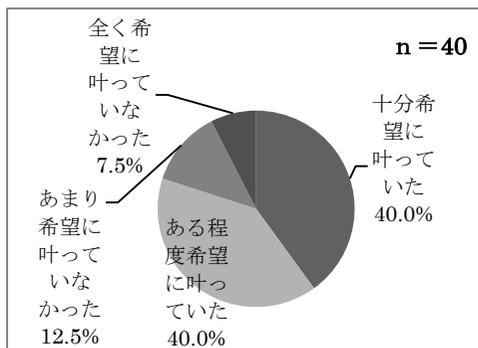


図2 ボランティア実習先への希望度合い

3) 実習中の行動についての自己評価と他者評価

上記のような目標と背景を抱えて参加した学生たちに、実習中の行動についての自己評価をしてもらった。結果は下記のとおりである（図3、表3）。この結果を見ると、「相手の話をよく聞く」ことや、「積極的に実習に参加したこと」について自ら高く評価（4.4）する一方で、「実習中に主体的に行動した」ことに対しては低い評価（3.3）を付けている。一所懸命取り組もうという意欲はあるものの、実際には自ら行動に移すことができないという、行動積極性（自主性、主体性）の弱さをここでも読み取ることができる。

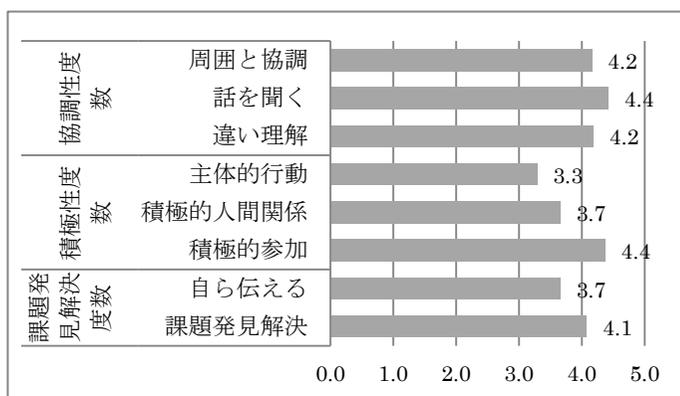


図3 実習先での行動についての学生自己評価

表 3 実習中の行動に対する自己評価

変数	自己評価項目							
	主体的行動	積極性 積極的 人間関係	積極的 参加	課題発 見解決	自ら伝 える	周囲と 協調	協調性 違いを 理解	話を聞 く
平均	3.2895	3.6579	4.3684	4.0789	3.6579	4.1579	4.1842	4.4211
標準偏差	1.2926	1.2793	0.9421	0.8505	1.2793	0.9733	0.8005	0.8584
不偏分散	1.6707	1.6366	0.8876	0.7233	1.6366	0.9474	0.6408	0.7368
最小値	1	1	2	2	1	2	2	1
最大値	5	5	5	5	5	5	5	5
中央値	4	4	5	4	4	4	4	5

評価項目のうち、「積極的参加」、「課題発見・解決」、「周囲と協調」の3項目については実習受入先担当者にも5件法で評価してもらうことを事前に依頼しており、学生本人にその結果が通知されるようになっていた。その評価（他者評価）と自己評価との差異（図4,5）から、評価の妥当性や公平性について学生たちがどのように捉えているかを考察した。

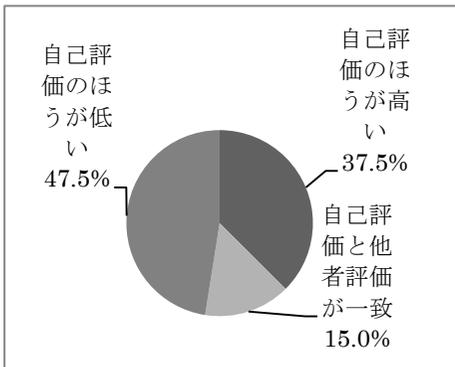


図 4 自己評価と他者評価の差異①

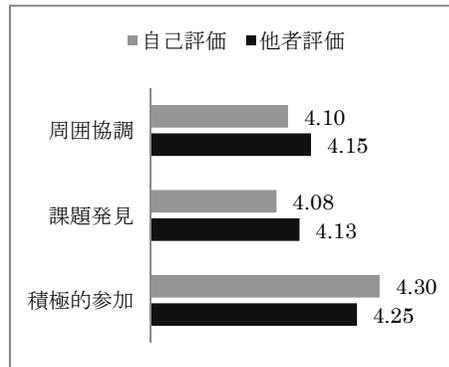


図 5 自己評価と他者評価の差異②

表 4 実習後の受け入れ先事業所からの評価（他者評価）

変 数	他者評価項目			
	積極的参加	課題発見解決	周囲と協調	他者評価計
n	40	40	40	40
平 均	4.2500	4.1250	4.1500	12.5250
標準偏差	1.1036	1.0424	1.0513	3.0549
不偏分散	1.2179	1.0865	1.1051	9.3327
最小値	1	2	2	5
最大値	5	5	5	15
中央値	5	4.5	5	14
母平均の 区間推定	信頼度	95%		
下限値	3.8970	3.7916	3.8138	11.5480
上限値	4.6030	4.4584	4.4862	13.5020

他者評価よりも自己評価が低い学生の割合は 47.5%である。自己評価と他者評価が一致している学生は 15.0%，自己評価のほうが高い学生は 37.5%であった。項目別にみると、「積極的参加」については、他者評価平均が 4.25、自己評価平均が 4.37 で自己評価がやや高い。「周囲との協調」では、他者評価平均が 4.15 に対して自己評価平均が 4.16、「課題発見・解決」については、他者評価が 4.13 に対して自己が 4.08 であり、他者評価のほうが若干高くなっている。

4) 評定の妥当性についての考察

次に、実習先から付与された評価について、その妥当性を学生がどう捉えているかを見てみよう。実習先希望度と評価の妥当性の関係を見ると、実習先希望度が高い学生は外部評価の妥当性を高く評価し、実習先希望度が低い学生は外部評価の妥当性を低く評価していることがわかり、統計的にも有為な結果となっている（表 5）。

また、自己評価の高低と他者（外部）評価についての妥当性の捉え方での関係を見ると、今回の調査においては「自己評価が低い学生は他者

(外部) 評価を妥当であるとみなす」傾向にあることも読み取れる (表 6)。これは, 自信のなさが自分の能力を低く見積もる傾向につながり, 加えて批判力の弱さから起こる現象でもあると考えられる。

表 5 実習先希望度の高低と評価の妥当性に対する捉え方

		評価の妥当性		全体
		妥当でない	妥当	
実習先希望度	4	1 14.3%	15 45.5%	16 40.0%
	3	1 14.3%	15 45.5%	16 40.0%
	2	3 42.9%	2 6.1%	5 12.5%
	1	2 28.6%	1 3.0%	3 7.5%
	全体	7 100.0%	33 100.0%	40 100.0%
独立性の検定			** : 1%有意 * : 5%有意	
カイ二乗値	自由度	P 値	判定	
14.0837	3	0.0028	**	

表 6 学生自己評価の高低と外部 (実習先) 評価の妥当性の捉え方

		外部評価の妥当性		全体
		妥当でない	妥当	
自己評価	高位	1 2.5%	5 12.5%	6 15.0%
	中位	4 10.0%	13 32.5%	17 42.5%
	低位	2 5.0%	15 37.5%	17 42.5%
全体	7 17.5%	33 82.5%	40 100.0%	
独立性の検定			** : 1%有意 * : 5%有意	
カイ二乗値	自由度	P 値	判定	
0.8183	2	0.6642		

ここでパフォーマンス評価の妥当性を考察するにあたって、佐々木ら(2006)の研究を参考にしたい。佐々木らの研究では、評定の正当性は評価者の特徴に影響されることに言及し、評価者の特徴を次の3タイプに分類している。

- ① 寛大化傾向（甘目に評価する。評価が尺度の上位に集中する）
- ② 中央化傾向（平均的に評価をつける。両極端の評価を避け中央に集中）
- ③ 厳格化傾向（厳しく評価する。評定が尺度の下位に集中）である。

また、これらの評価者による評定のバラツキが、評価の信頼性にどのような影響を及ぼすかを、一般化可能性係数(Brennan, 2001)を用いて分析し、一定の結果を導き出している。彼らの調査では、評価者の「評価にバラツキが大きい」ほうが信頼性が高く、「評価のバラツキが少ない」場合に信頼性が低くなるというのである。この結果から佐々木らは、パフォーマンス評価を行う際には、評定のバラツキが大きい評価者(同点に集中させない評価者)を評価者集団に含めることを奨めており、キャリア教育における評価システム構築の一つの示唆となっている。

今回の実習先からの評価に目を向けると、評価を依頼した3項目全てに5点(満点)を配したものが21件で、全体の48%を占めている。平均でも4.25ポイントであり、かなり高い評価、つまり甘い評価がつけられたことがわかる。これは、学生たちの行動が優れていたと素直に見ることもできようが、わずか4日の出来事で、しかも限られた仕事・活動を見て評価することの困難さを考えれば、好意的かつ無難に回答をしたということも推測できる。特に、「課題を見つけ、解決していたか」という設問は、学生の感想に指摘があるように、回答の難易度が高い項目であったと思われる。

これらの結果を上述した佐々木らの理論に当てはめて考えるならば、

今回の実習先からの 5 件法による評価では、かなり集中的に高評価が与えられていることから、パフォーマンス評価としての信頼度は高いとは言えないことになる。それを裏づけるかのように、評価シートの欄外に設けた自由記述欄には、5 件法評価からは読み取れなかった厳しい評価が数多く書き込まれていた。コメントに書かれた評価に関するキーワードをポジティブ・ワードとネガティブ・ワードに分類したものが(表 7)である。頻出語を抽出すると、「積極的に行動してくれた」、「提案してくれた」、「時間を守って動いていた」、「真面目で素直に取り組んでいた」ことが好意的に評価される一方で、「積極性は見られなかった」、「規律性を身に付けてほしい」、「一般常識・マナーに欠けていた」という厳しい指摘が、ポジティブ・ワードの記述数を超えていたのである。

このように自由記述されたものと照らし合わせて分析してみると、5 件法による評価の点数のみでは評価者が実際に伝えたいことを伝えきれていないことが窺える。一見、客観性がありそうな数字も、パフォーマンス評価である以上、評価者の心理に左右されることは避けられず、妥当性は高くないことが推測できる。実習先事業所の方々が、日頃大学生と接する機会が少なく、評価活動に携わる経験が少ないとすれば、短期間学生と接触して成績評価に携わることには、それなりの心理的負荷をかけていたように思われる。そのため、無難なアクションとして満点をつける、あるいは、バラツキのないポイントを付けるという結果になったと考えられる。事業所間での評価格差が少なく、また評価する人の負担を軽減する手法を今後検討していく必要があるだろう。

次に、学生たちは実習先からもらった評価についてどう思っているのかを見ていく。自分に付けられた評価が「妥当であると思う」と回答した学生の自由記述コメント(表 8)では、「積極的・意欲的取組みを評価してもらえた」、「指摘事項は自分でもそう思うので妥当。よく観察

表7 ボランティア受入先担当者からの記述評価コメント（分類）

	ポジティブ・ワード	ネガティブ・ワード
積極性・主体性に関連するもの	<ul style="list-style-type: none"> ・積極的（4）・提案していた（2） ・自分から見つけて行動 ・意欲的・リーダーシップ・物怖じせず・率先して・個性的・目的を持ち・目標を持ち・若々しい発想力 	<ul style="list-style-type: none"> ・積極性に欠けていた・積極的に会話をしてほしい・積極的な発見や提言をしてほしい・積極的に活動したほうがよい・積極性はみられなかった・意欲が見られなかった ・目的、課題を持ってほしい ・課題を見つけ出すまでではできていない・動作がきびきびしていない
規律性に関わるもの	<ul style="list-style-type: none"> ・時間を守っていた（3） ・時間どおり 	<ul style="list-style-type: none"> ・遅刻が残念（2）・規律性を・勝手に行動・自分の中だけで解決していた・協調性を意識しよう
態度・常識・マナーに関わるもの	<ul style="list-style-type: none"> ・真面目（3）・素直（3）・話をよく聞いて行動・的確に活動・正しい理解と伝達・献身的・はきはき・明るく・効率よく・優しく・しっかりとした言葉で挨拶・指示にしたがい行動・耳を傾け・熱心・快く引き受け・誠実・能力の高さ・周囲への気配り・すぐに確認・嫌な顔ひとつせず・謙虚な気持ち・自分の言葉で堂々・協力・気づき 	<ul style="list-style-type: none"> ・一般常識が不足・緊張している様子・声が小さい・やる気のなさが気になった・立ち方、挨拶や言葉遣いに気を付けて・靴を揃えて・もっと笑顔を・身だしなみに配慮を・言葉遣いや姿勢、礼儀正しい態度（挨拶）を・質問、挨拶・返事ができない・手先不器用

してくれている」といったものである。一方、評価が「妥当ではない」と回答した学生のコメントは、「自分としては頑張ったつもりだが、評価が厳しすぎる」、「自分の中ではいろいろ反省点があるのに良い評価だった。逆に不安になる」、「オール5（満点）の評価は少し良すぎると思う」というものがある。全体に甘い評価の中で、数件の事業所においては極端に低い評価を付けていた所もある。その事業所で実習した学生のうちの数人は、この評価の妥当性について疑問を持っていた事実があり、こうしたケースをどのように捉えるべきかの検証が必要である。

最終的な単位認定は、学生の自己評価と実習先評価結果を踏まえた総合的な評価を付与することになるが（図6）、実習先事業所評価、学生自己評価ともに数値には現れない「マイナス点」あるいは「成長度」をどのように扱っていくかが課題である。

表8 実習先評価について「妥当」「妥当ではない」と思う理由

評価が妥当と思う理由	評価が妥当でないと思う理由
<ul style="list-style-type: none"> ・積極的に活動し、そこを評価していただいた。 ・意欲的に取り組んだところを評価してもらえた。 ・コメントに書かれていたことは、その通りだと思ったので。 ・マンツーマンで働きを見てくださっていたので、評価は妥当だと思う。 ・周囲との協調性について指摘されたが、自分でもその点は欠けていたと思うので妥当だと思う。 ・よく観察してコメントをしてくださったと思う。 ・個人的にはうまくできなかった部分があるが、努力や取り組みの姿勢を見て評価してくださった。褒めてもらったことが嬉しい。 ・自ら進んで行動するということで低い評価だった。自分もそう感じていたので妥当である。 ・他の2人より積極性が足りなかったという評価を受けた。その通りである。 	<ul style="list-style-type: none"> ・作業内容が単純だったのでどうなのかと思う。（評価は良かったが） ・自分の中ではいろいろ反省点があるのに、良い評価だった。 ・オール5（満点）の評価は少し良すぎると思う。 ・評価の仕方が厳しいと思う。 ・評価が全部5だった。しかし、本当に悪いところがないのか、心配になった。 ・自分としては頑張ったつもりであるが、ダメだしされた。評価は厳しい。 ・初めての経験で緊張していたことは確かだが、自分としては精いっぱいやったつもりである。だが、ついてきた評価は最低のものだった。納得がいかない。だが、振り返りをしたり、友達や先生たちと話したりするうちに、実はいちばん大変な思いをした自分たちが一番勉強になり、成長できたのではないかと思えるようになった。 ・受入先からは、ほとんど説明や指示がなく、何をしてよいのかわからないまま自分で動くしかなかった。一緒に働いている人からも声をかけてもらうこともなく終わってしまった。積極性がないと厳しい評価を受けたが、そういう評価が正しいのかどうかかわからない。 ・仕事をしているところを誰も見ていないのに、誰がどう評価をつけたのか疑問である。 ・あの仕事内容で評価をつけるほうが難しいと思う。

科目担当者による単位認定（包括的評価）： 出欠状況と課題提出+学生・実習先評価結果をもとに、能力を推測して評価する
実習先担当者（他者評価）： 実習中の行動・能力そのものを評価する
学生（自己評価）： 自分の意欲、行動、能力について省察し評価する

図6 実習評価の構造：評定者と評価の対象

まとめと今後の課題

大学の質保証が問われる時代にあって、大学の出口と社会の入り口を繋ぐキャリア教育への期待は確実に高まりつつある。本研究のため、北海道内外の大学のキャリア教育の現状を見てきたが、2011年の法令化以降のキャリア教育導入と定着のスピードは目ざましいものがあり、キャリア教育が第2ステージに突入したことが認められた。第2ステージにおける課題は、まさに教育内容の質の確保と学士課程内における体系化および効果測定である。何よりも内容の精査充実と教育水準の維持なしには、大学教育の質保証の一翼を担うことには繋がらない。自立した社会人、一市民として働き貢献する意識を醸成し、実社会で通用する水準の知識と技能の修得を目指すプログラムの整備が必要である。それは学士課程内の他科目と切り離されたものではなく、カリキュラム全体が体系化されたものであることが望ましい。そのためにはカリキュラム・マップを整備し育成プロセスと到達目標を可視化することが先決であろう。

また、キャリア教育科目の効果測定はパフォーマンス評価や真正評価が主流であることから、公平性や妥当性についての問題が生じることは避けられない。「対象の観察」によって評定をつけるパフォーマンス評価には評価者の主観が混入しやすく、人によって結果が異なることも考えられるため、信頼性と妥当性について常にケアが必要である。

本研究では、ボランティア実習における学生の自己評価と実習先評価を考察することで、パフォーマンス評価の妥当性について検証した。実習先からの評価が極めて低かった学生群は、評価の妥当性を低く捉えており、学生調査の記述欄には、「自分なりに精一杯前向きに取り組んだが、少しも評価されず納得がいかない」等と書いているものもあった。ところが、その後、筆者が面接法を用いて振り返りを行わせたところ、「振り返りをしたり、友達や先生たちと話したりするうちに、実は誰よりも大変な思いをした自分たちが一番勉強になり、一番成長できたのではないかなと思えるようになった」という言葉が飛び出してきた。付与された評価が学生の省察と成長に繋がった瞬間であった。他方、実習先から満点の評価を得た学生の中には、「自分にはいろいろ反省点があるのにとっても良い評価だった」、「オール5だったが、本当に悪い点はないのか心配になった」というコメントもあり、パフォーマンス評価の難しさが顕わになっている。

これら評価の妥当性課題を克服する方策の一つとして、評価ツールであるルーブリックの活用について検討した。事前に評価規準と基準を明らかにすることで評価の曖昧性を抑制することができることとされるルーブリックは、評価者の負担を軽減することにつながる。また学修者にとっても自分の達成度合いや長所の発見に繋がる効用があり、評価者・被評価者双方にとってのメリットが認められ、キャリア教育科目の採用可能性が高いものであることがわかった。

インディアナ州立大学の George D Kuh は NSSE という大学生調査開発の際に、“Student Engagement ”という言葉を用いている (George D Kuh, 2001)。これは学生が大学という場にどれだけ主体的に関わり成長できたかを測るものであるが、それはまた同時にそのための環境と教育の質を大学はどれだけ提供できたかを問い直すものでもあろう。知識の獲得

だけを目的としないキャリア教育の評価は、その学生の大学生活全体を評価することでもあり、学生の変容と成長を促すものでなければ意味がない。学習の過程、改善・発展の有り様を抜きにして正当な評価にはならないであろう。評価の公平化・精緻化のためには、手間はかかるが、最終的には個別面接法の導入が不可欠であると考えている。さらに学修者の自律性を重視して学生本人の自己評価も多用するため、評価者である教員と評価を受ける学生が、共に評価の在り方や考え方を学ぶことも必要になるのではないか。

キャリア教育評価の目的は何か。誰のための、何のための評価なのか、「成長する存在である学生」よりも、「評価する自分（教員）」のものになってはいないか、この研究を通して奇しくも自省の機会を得ることとなった。日本の大学におけるキャリア教育の評価に関する研究は緒についたばかりである。今後は先進的な海外の事例にも目を向け、科学的な論証に繋げていきたい。

[註]

- (1) 文部科学省 キャリア教育の推進に関する総合的調査研究協力者会議 報告書 (2006), 中央教育審議会答申「今後の学校におけるキャリア教育・職業教育の在り方について」(2011) 参照
- (2) アスティンのカレッジ・インパクト: 大学という環境が学生の学修経験や情緒面での自己認識に影響を及ぼし、成果と成長につながることをアスティンは I(input)-E(environment)-O(output)モデルを使って示した。
Astin, Alexandar.W.1993. *Assesment for Excellence : The Philosophy and Practice of Assesment and Evaluation in Higher Education*, Phenix, Amazona: ORYX Press. p45
- (3) 関西国際大学「KUIS 学習ベンチマーク」のループリック: 吉田武大 (2011) 「アメリカにおけるバリューループリックの動向」『関西国際大学教育総合研究 所研究叢書 4号』より抜粋

- (4) 今回事例としてとりあげた「ボランティア II」は筆者が担当している科目ではないが、ボランティアを広義のキャリア科目として捉え、科目担当責任者よりデータの活用について許諾を得ていることを付記しておく。

[引用・参考文献]

- 高浦勝義 2000 『ポートフォリオ評価入門』 明治図書
森敏彰・秋田喜代美 2000 『教育評価 有用用語 300 の基礎知識』 明治図書
山田礼子 2012 『学士課程教育の質保証へむけて 学生調査と初年次教育からみえてきたもの』 東信堂
橋本重治 2003 『教育評価法概説 2003年改訂版』 図書文化
佐々木亮 2010 『評価論理 評価学の基礎』 多賀出版
高浦勝義 2004 『絶対評価とルーブリックの理論と実際』 黎明書房
上西充子 2007 『大学のキャリア支援-実践事例と省察-』 経営書院
田中耕治 2012 『パフォーマンス評価入門』 ミネルヴァ書房
田中耕治 2004 『学力評価の“今”を読みとく』 日本標準
田中耕治 2008 『教育評価』 岩波書店
久保田賢一・岸磨貴子 2012 『大学教育をデザインする-構造主義に基づいた教育実践』 晃洋書房
梶田叡一 2010 『教育評価 第2版』 有斐閣双書
土持ゲーリー法一 2010 『ポートフォリオが日本の大学を変える』 東信堂
国立教育政策研究所監訳 2010 『PISA2009年調査 評価の枠組み OECD生徒の学習到達度調査』 明石書店
文部科学省 2012 『諸外国の教育動向 2011年度版』 明石書店
石井 英真 2012 「新しい評価の動向：思考力・判断力・表現力を育てるパフォーマンス評価（特集 教育における評価の現状と課題）」『京都大学教育学部教育展望』58(7), pp36-40
上西充子・川喜多喬 2010 『就職活動から一人前の組織人まで 初期キャリアの事例研究』 同友館
刈谷剛彦 2012 『アメリカの大学・ニッポンの大学 TA, シラバス, 授業評価』 中央公論新書
寺嶋和夫 2009 「キャリア教育の有効性と方向性に関する実証的研究(1) 龍谷大学経営学部キャリア教育の試みと「実践・キャリア形成論 I」受講生の評価を中心に」『龍谷大学経営学論集』49 (1)
Alverno College Faculty, *Ability-Based Learning Outcomes(6th edition)*, Alverno College Institute, 2005

- 松下佳代 2007 『パフォーマンス評価-子供の思考と表現を評価する』日本標準ブックレット No. 7, P. 6
- 黒上晴夫 2005 「Rubric 評価の考え方と活用」『IMETS』158号(財)才能開発教育研究財団・教育工学研究協議会, pp50-55
- 富永美佐子 2008 「進路選択自己効力に関する研究の現状と課題」『キャリア教育研究』25(2), pp97-111
- 内浦有美, 毛受芳高 2008 「キャリア教育の評価-情動の喚起と気づき・意欲・行動の変容の関係性に着目した評価視点の提唱」『Works Review3』pp196-209
- 佐々木典彰, 村木英治 2006 「評定者内の評定のばらつきが信頼性に及ぼす影響」『教育情報学研究』第4号
- 黒上晴夫 見える「評価」で授業が変わる～ルーブリックで授業作り
http://www.justsystems.com/jp/school/academy/hint/rubric/ru01_01.html
(2013. 1. 10 取得)

(わだ よしこ, 札幌大谷大学社会学部教授)