

パフォーマンス評価におけるポスト実証主義的アプローチの可能性

北 川 剛 司 奈良教育大学大学院 (教職開発専攻)

(平成25年5月7日受理)

A Study for the Possibility of Post-Positivist Approach in Performance Assessment

KITAGAWA Takeshi

(School of Professional Development in Education, Nara University of Education)

(Received May 7, 2013)

Abstract

Currently, in the method of performance evaluation in Japan, positivistic approach that assesses the target by using the evaluation criteria such as rubrics have become mainstream. But, Hart introduced not only the positivistic method but also the non-positivistic method in performance assessment.

Now, the positivistic approach is positioned as the meaningful method of classroom assessment. But, there are also a number of criticisms against this situation. According to the Joint Committee, there are four standards to judge the appropriateness of the evaluation: utility standards, feasibility standards, propriety standards, and accuracy standards. However, much of criticisms in today's discussion are those from the accuracy standards.

There are some aspects that cannot be assessed by a positivistic approach. So, we need to study from the perspective of the post-positivist approach. Art based research as a post-positivistic approach, which was proposed by Baron and Eisner, given an indication about the possibility of performance evaluation based on the post-positivistic approach.

キーワード：パフォーマンス評価、ポスト実証主義、
アート・ベースド・リサーチ

Key Words：Performance Assessment, Post-Positivism,
Art Based Research (ABR)

1. はじめに

従来のペーパーテストによる正誤問題・多肢選択問題・短答問題で測ることのできる学力の限界が指摘され、もっと広い意味での学力を測るべく「パフォーマンス評価」に注目が集まるようになって久しい。松下佳代によると、パフォーマンス評価とは、「ある特定の文脈のもとで、様々な知識や技能などを用いて行われる人のふるまいや作品を、直接的に評価する方法」⁽¹⁾であると述べられる。また、その方法については、『パフォーマンス課題 (performance task)』を与えて解決・遂行させ、それを複数の評価者が、『ルーブリック (rubric)』と呼ばれる評価基準表を用いながら、評価していきます⁽²⁾とされる。すなわち、予め用意したルーブリック (評

価基準表) を使って、観察された対象のパフォーマンスを解釈し、学力を把握しようとするのが松下のいうパフォーマンス評価の方法である。このような評価は、実証主義⁽³⁾ (positivism) の立場をとる評価 (実証主義的アプローチにもとづく評価) ということができる。

本研究でいうところのポスト実証主義的アプローチとは、アメリカにおいてタイラー (Tyler, R. W.) が教育プログラムの評価において用いた実証主義的な方法の後、主に1960年代以降に提唱された、教育プログラムの評価の多くが用いた、実証主義に依らない評価方法 (オルタナティブ評価) の総称である。

本研究の目的は、パフォーマンス評価におけるポスト実証主義的アプローチの可能性と意義を明らかにすることである。

2. パフォーマンス評価の現在—ハート(Hart, Diane)を手がかりに—

ハートは全米的によく知られた評価コンサルタントであり、わが国においてはハートの著書『Authentic Assessment: A Handbook for Educator』が田中耕治氏の監訳で2012年に邦訳されている(邦題『パフォーマンス評価入門—「真正の評価」論からの提案—』)。本章ではこの著作を手がかりにパフォーマンス評価(論)の現在について整理したい。

我々は、パフォーマンス評価と聞くと、パフォーマンス・テストを想起することが多い。だが、そのような理解はパフォーマンス評価の一面的なものにすぎない。実際は、「テストは多々ある評価形態のなかのただ1つの形態にすぎない」⁽⁴⁾と言われるように、パフォーマンス評価の方法はテスト法に限られない。ハートはパフォーマンス評価の様々な方法・形態を、それぞれが提供する生徒についての情報の種類にもとづいて3つの大きな種類に分類している。すなわち、①観察(observations)、②パフォーマンス事例(performance samples)、③テスト(tests)である。

2.1. 観察にもとづく評価の方法

「教師は手にクリップボードを持って教室のなかを歩き回り、定期的に生徒についてのメモを取ります。これが観察法の本質です」⁽⁵⁾といわれている。このようにして得られる日常的な観察記録は、子どもの姿を可視化された情報に変え、実践の振り返りや達成の点検、および、同僚との情報共有において活用されるなど、まさに教育実践の一環として位置づいている。だが、観察を評価の方法として位置づける場合は、そうした日常的な観察法とは区別される必要がある。すなわち、評価としての観察は、教師の非公式なメモとは区別され、体系的な評価プログラムとならなければならない。そのために、評価としての観察は、次の3つのことに取り組まねばならない⁽⁶⁾。

第一に、観察を体系的なものにすることである。そのためのガイドラインは次の通りである。

- ・全ての生徒たちを観察する。
- ・しばしば、そして定期的に観察する。
- ・観察結果を書いて記録する。
- ・非典型的なものも典型的なものも書きとめる。日常の観察は、いつもと違うことを観察する場合と全く同じくらい価値がある。
- ・信頼性を高めるために多様な観察結果を集める。1つの例だけで、それをパターン化することはできない。
- ・観察の妥当性を高めるために、さまざまな文脈での

証拠を統合する。

第二に、観察の対象を明確にすることである。つまり、観察において何に目を向けるべきかを明確にすることである。そのために役立つツールが「発達チェックリスト(developmental checklist)」と「インタビュー用紙(interview sheet)」である。これらのツールはアメリカの教室においても最も一般的に用いられている。

前者には、発達にしたがって評価すべき学習行動の特徴(traits)が記述されており、教師はそれを用いて時の経過のなかで生徒の進歩を明確にみとめることができる⁽⁷⁾。後者は、一般的には、教師が生徒に尋ねようと計画している質問のリストと質問への応答を記録するスペースで構成され、教師はリストにあがった質問をしながら、応答を記録していく⁽⁸⁾。

第三に、観察したものを記録に残すことである。観察法が信用できるとみなされるためには、記録に残すことが重要である。観察したことを記録するには膨大な時間がかかることがしばしば問題となるが、書くことや記録を取ることに使う時間を最小限にするメモ取りや記録保存のシステムがつくられることによって、そうした問題は軽減しうる⁽⁹⁾。例えばコンピューターのデータベースや宛名ラベルの活用などがそれにあたる。アメリカにおいてはすでに多くの教師によってこのような記録保存システムが活用されている⁽¹⁰⁾。

2.2. パフォーマンス事例の収集にもとづく評価の方法(ポートフォリオ評価法)

「ポートフォリオは個人のスキルやアイディア、関心、成果を示す証拠が入っている容器」⁽¹¹⁾である。容器は、ファイルやノートであったり、LD、CD、DVDといったハイテクのものであったりさまざまである。ポートフォリオは、特定の時間と場所における生徒の達成の断片を提供するテストとは違って、時の経過のなかで学習を記録するという特徴をもつ⁽¹²⁾。ポートフォリオの内容の中核となるのは、生徒の作品の事例および自分たちの作品に対する生徒たちの振り返りである⁽¹³⁾。

ポートフォリオには生徒の学びの履歴が継続的に蓄積されていることから、ポートフォリオを評価対象として中・長期的な指導や学習のあり方を評価することが可能である。ポートフォリオの評価では、スタンダードを使った方法のほか、討論、批評、コメント法などを用いる方法が紹介されている⁽¹⁴⁾。

2.3. パフォーマンス・テスト(パフォーマンス課題)にもとづく評価の方法

ポートフォリオ評価法では、時間の流れのなかで継続的に作成されたファイル等が評価の対象となるのに対して、パフォーマンス・テスト(パフォーマンス評価)は

決められた時間と場所において、パフォーマンス課題 (performance tasks) を与え、その課題への取り組みにおいて発揮された生徒の力量を評価しようとするものである。パフォーマンス課題を「現実の世界からの挑戦や問題を模した課題」⁽¹⁵⁾とすることで、そのパフォーマンス・テストは「真正の評価 (authentic assessment)」となる。

形式と利用という視点から、ハートはパフォーマンス課題を三つに分類している⁽¹⁶⁾。すなわち、「短い評価課題 (short assessment tasks)」⁽¹⁷⁾、「より大掛かりなイベント課題 (event tasks)」⁽¹⁸⁾、「長期にわたる拡張課題 (extended tasks)」⁽¹⁹⁾である。

「短い評価課題」は、生徒がある学習内容のなかで、基本的概念、手続き、関係、思考スキルなどをどれほど習得しているかを判断するために、「オープン・エンドな課題 (open-ended tasks)」への取り組みや「発展的な多肢選択式問題 (enhanced multiple-choice)」や「概念地図法 (concept map)」によって、評価を行うものである。

「イベント課題」は、生徒がすでに知っていることだけでなく、さらには彼らがどれほど上手くその知識を使用しているかを浮き彫りに出来るように作成される。その評価は、「教師の観察、生徒のパフォーマンス事例 (performance samples) の採点、自己評価、仲間による評価、またこれらのアプローチを一部組み合わせたものに基づいて実施されます」⁽²⁰⁾とされる。

「拡張課題」は長期にわたる多目的プロジェクトで、学期や学習単元の最初に割り当てられる。その評価は、口頭や書面での応答、小論文やプレゼンテーションなど多岐にわたって行われる。この課題にもとづく評価は「結果を簡単に採点したり報告したりするには役立ちません」⁽²¹⁾とされるように、評価のために多くのコストを要する。だが、実際の生活において応用可能な能力を評価することができる点など、その意義は認められており、利用が続けられている。

以上、ハートにおけるパフォーマンス評価の方法においては、採点ルーブリックを用いた実証主義的な方法以外にも、観察にもとづく評価や、ポートフォリオ評価を介した討論、批評、コメント法、および概念地図法など、実証主義的でない方法も同時に用いられていることが分かる。すなわち、パフォーマンス評価は、実証主義的アプローチと実証主義に依らないアプローチの両側面からその方法が検討されている。

だが、『「真正の評価」論は、『目標に準拠した評価』をパワー・アップするために、提起されたものです』⁽²²⁾という記述に見られるように、わが国では、パフォーマンス評価の一種である真正の評価が、目標標準評価の延長線上にしばしば位置づけられてきた。こうしてわが国

においては、パフォーマンス評価でルーブリックを用いることが一般的であると理解されていくこととなっている。

3. 評価における実証主義的アプローチの意義

3.1. 相対評価から到達度評価へ

相対評価の非教育的な性格および評価としての妥当性の不十分さが自覚され⁽²³⁾、相対評価に代わる新たな評価方法として1970年代半ば頃から到達度評価に注目が集まった⁽²⁴⁾。

それにもかかわらず、相対評価は70年代半ば以降もわが国では多くの学校で引き続き用いられたが⁽²⁵⁾、相対評価に対する根気強い批判や到達度評価の研究と実践の蓄積が実を結び、21世紀を前後する頃より到達度評価(目標標準評価)が多くの学校において全面的に採用されるようになっていった⁽²⁶⁾。

3.2. 目標標準評価の意義

先にも述べたように、目標標準評価はその方法論から考えて、実証主義的アプローチにもとづく評価であるといえる⁽²⁷⁾。そこで、目標標準評価の意義について整理することで実証主義的アプローチの意義への示唆を得たい。

目標標準評価の意義は次のように整理できる。第一に、『「目標に準拠した評価」では、まさしく学力の中身である『目標』にどの程度到達しているかが明確となる』⁽²⁸⁾こと。すなわち、教師ないし学習者が、評価の時点において、教育(学習)目標をどれだけ達成できたかということが明確になるという意義である。

第二に、「この公共社会を生きるのに必要な学力をすべての子どもたちが身につけることが必要である」⁽²⁹⁾という学力保障論の立場から、つまりいた子どもに対する回復学習が可能となること。すなわち、子どもがどこでどのようにつまづいているのかということ、および、今後の到達基準が分かるのでそれを補う手立てを講じ易いという意義である。

第三に、排他的競争から解放され共通の目標に到達することが目指され、学習における協同を生み出すこと⁽³⁰⁾。すなわち、励まし学び合うことを推し進める教育評価であり、子どもたちが共通の目標に向かって共に成長しようとする教育を生み出し易くなるという意義である。

その他、目標標準評価の意義に関わって個人内評価の意義がある。個人内評価とは、目標標準評価の一種で、規準ごとに全員共通の基準を用いるのではなく、子ども個人ごとの基準によって評価する方法である。その意義は、『「個人内評価」を合わせて用いることで、一人一人の子どもたちの良い点や可能性、進歩の状況を把握する

ことができる」⁽³¹⁾とされるように、共通の到達基準を用いるよりも、より子ども一人一人に寄り添った形で達成(到達)が明確になるという意義である。

さらに、目標標準評価の必須過程である規準と基準づくりそのものにおける意義がある。すなわち、その分野の達成についての評価研究自体が、カリキュラムや教材の研究となるという意義である。

4. 評価における実証主義的アプローチに対する批判

4.1. 評価の優良基準

評価に関する合同委員会(the Joint Committee on Standards for Educational Evaluation)は、評価の適切性を問い直す(評価の評価、メタ評価)ための優良基準を示している。その優良基準とは、①有用性(utility)、②実現可能性(feasibility)、③正当性(propriety)、④正確さ(accuracy)の4つの領域(category)である⁽³²⁾。

有用性基準とは、「評価が有益で、適時で、影響力のあるものになっているかどうかを判断するものである」⁽³³⁾。すなわち、評価の受け手としての役割を受け持つ人(オーディエンス)が必要としている実践的な情報に、その評価が果たして役立っているのかどうかを定義するのが有用性基準である。有用性基準を運用するためには、評価者自身が自分たちの評価の受け手は誰であるのかを明確にし、彼らが何を必要としているのかを確かめることが必要である。

実行可能性基準は、「評価が現実的であり、俊約的で、駆け引きがうまくて、経済的であるかどうかを判断するものである」⁽³⁴⁾。評価は通常、実験室ではなく、自然な状態の中で執り行われるものであること、および評価は価値のある資源を消費するものであることは広く知られている。それゆえ、評価の設計はフィールドの状態において運用可能なものでなければならないし、評価は多くの資源や、材料や、人材や、時間を必要以上に消費するものであってはならない。

正当性基準は、評価が実に様々な人に影響を与えるという事実によく目を向ける。ゆえに、正当性基準の利用が意図するのは、評価によって影響を受ける個人の権利の保護を促すことである。正当性基準は、「評価を行う者が、不法で、無節操で、非倫理的で、不適当な振る舞いを行っていないかどうかを判断する」⁽³⁵⁾。

正確さ基準は、「評価が健全な(sound)情報を産出したかどうかを判断するものである」⁽³⁶⁾。すなわち、正確さ基準とは、とらえたいものをその評価によってとらえられたかどうかを判断するものである。評価によって得られる情報は、技術的に妥当で、評価によって言い渡された判断は論理的にデータと結び付けられていなくて

はならない。

4.2. オルタナティブからの批判

相対評価でなされた批判を乗り越える評価方法として、実証主義的アプローチにもとづく目標標準評価が注目を集める一方、目標標準評価についてもそのオルタナティブ⁽³⁷⁾の立場からさまざまな問題性が指摘されてきた。

(1) 評価規準(基準)自体の妥当性の問題

到達目標として設定した評価の規準や基準自体がそもそも対象の意図した側面をとらえるのに妥当でない場合があり得るという批判である。ただし、これは目標標準評価の本質的な問題というより、目標標準評価を適切に行えていない場合に指摘される問題である。すなわち、適切な到達目標(評価規準)づくりを意識して行くことで解消できる可能性のある問題である。これは、妥当性のある規準づくりという技術の問題であることから、正確さ基準による批判の一種といえる。

(2) 評価の規準以外の学びの側面を捉え損ねる可能性があるという問題⁽³⁸⁾

到達目標を規準としてそれが子どもたちに獲得されたかどうかという視点でのみ授業をみていくと、教師のねらいからはみ出したり、それを乗り越えていく子どもたちの姿が軽視されたり無視されることが起こるという批判である。これは、学びの側面全体を評価規準でカバーできていないという技術の問題であることから、正確さ基準による批判の一種といえる。

(3) 子どもたち自身による内的な評価を十分に位置づけていない(指示待ち、評価待ちの子どもを育てる)⁽³⁹⁾

自分たちの活動が教師のねらいとしての規準にもとづいて常に評価されると、子どもたちはそれを意識せざるをえない。こうして教師の意図(教師がどうしてほしいのか)を極端に意識し、自分たちはどう思うのかということが軽視されていくという批判である。これは、評価の結果が誤った形で活用され、子どもたちの学びの可能性をかえって狭めてしまうという評価結果の運用の問題であることから、有用性基準による批判の一種といえる。

(4) 目標標準評価は結果主義の評価である(プロセスを評価できない)⁽⁴⁰⁾

到達目標に達したかどうかを問題にするあまりに、目標到達に至るまでの過程における子どもたちの努力、試行錯誤、つまづき、葛藤を不問にするという批判である。これは、学びのプロセスを評価規準でカバーできていないという技術の問題であることから、正確さ基準による批判の一種といえる。

(5) 量的な評価にとらわれて高度な学力を対象とする質的な評価への目配りが弱くなる⁽⁴¹⁾

目標標準評価では、その評価規準である到達目標に

よって対象をとらえようとするものであり、ゆえに、規準化可能で観察可能・伝達可能な範囲の量的な側面にしか着目できず、それよりも高度な質的な側面をとらえていないという批判である⁽⁴²⁾。これは、学びの質的な側面を評価規準でカバーできていないという技術の問題であることから、正確さ基準による批判の一種といえる。(6)目標準拠評価には多くの時間的・費用的コストがかかる

目標準拠評価が必要とする評価基準表の作成にはコストがかかる。評価基準表の作成には次のようなプロセスが必要となる。①評価基準表は評価によってとらえたいものを予め決定し、規準化する。②さらに、それぞれの規準ごとに達成の基準を設ける。③基準の信頼性を確保するために、評価者間で基準の調整（モデレーション）を行って基準の統一化を図る。以上のプロセスをふまえて、評価基準表を作成するには、かなりの時間的・費用的コストがかかる。このようなプロセスを含む評価を実践の中で日常的に活用するにはコスト面の困難がある。これは、日常的な評価をすべて目標準拠評価で行うことの不可能性を指摘するものであり、実行可能性による批判の一種であるといえる。

以上、オルタナティブからの批判を整理すると、(1)(2)(4)(5)は正確さ基準による批判である。(3)は有用性基準による批判である。(6)は実行可能性基準による批判である。すなわち、正確さ基準による批判が多い一方で、取り上げた批判のなかには、正当性基準による批判はなかった。だが、評価結果が個人の権利に及ぼす影響などが検討されなくてはならない⁽⁴³⁾。今後、正当性基準の面からも問い直しが求められるだろう。

5. ポスト実証主義的アプローチにもとづくパフォーマンス評価—バロン (Barone, T.) とアイズナー (Eisner, E.W.) のArts Based Researchを手がかりに—

5.1. バロンとアイズナーについて

バロンは、現在アリゾナ州立大学の教授で、30年以上前に、スタンフォード大学において、教育問題についてのリサーチや記述における文芸的随筆（ノンフィクション）の可能性について調査し博士号を取得している人物である。

アイズナーは、現在スタンフォード大学の教授である。アイズナーは1960年代以降よりポスト実証主義の立場からの評価論を展開してきた人物である。アイズナーによる、教育的鑑識眼や芸術批評といったオルタナティブ評価の提唱はわが国のカリキュラム構成論、教育評価論においても広く認知され影響を与えてきた。

本章ではポスト実証主義的アプローチ研究の最新の動

向として、バロンとアイズナーの述べるアート・ベースド・リサーチ (Arts Based Research) (以下、A B R という) に着目し、それを主な手がかりとして、ポスト実証主義的アプローチにもとづくパフォーマンス評価の方法について考察し、パフォーマンス評価におけるポスト実証主義的アプローチの可能性についての示唆を得たい。

5.2. A B R 提唱の背景

バロンとアイズナーがA B Rを提唱した背景にはリサーチにおける自然科学的方法への万能視に対する批判意識があった。

我々は、リサーチというともっぱら公式化され、練り上げられたものとして考えがちである。そこではあいまいな部分を減らすことこそが主な美徳としてみなされる。我々が誤って正確さだと理解しているものにひたすら没頭することは、リサーチの方法論の標準化や、自然科学の手続きこそが模範とすべき手続きであるとみなすことにつながる。我々はこうした自然科学の方法で物事の全てをとらえたように思っていることが多い。だが、実際は、そうしたリサーチの方法は、ものごとの全容 (whole story) をとらえることとは程遠いのである⁽⁴⁴⁾。

すなわち、バロンとアイズナーにとってのA B Rとは、自然科学的方法がとらえたと思っていることが限定的であって、そうした方法ではものごとの全容はとらえられないということに対する問題意識のもと提唱されたオルタナティブとしての評価方法なのである。

5.3. A B R の特徴

まずはA B Rの主体について。A B Rを行えるのは、アーティスト (artist) である。アーティストとはプロの (professional) 画家や詩人やダンサーや小説家を必ずしも意味しない。だが、誰もがA B Rを行えるわけではない。A B Rを行うアーティストには、芸術家としての手腕 (artistry) が求められる。芸術家としての手腕を持ったアーティストこそが、「心で感じたり見たりするのと同様に全身を使って探究したりテストをおこない、次に、応答したり再度テストを行い、それから複雑な様子を記録し、最終的に、それらを分類したりパターン化したり意味を産出する」⁽⁴⁵⁾ことのできるA B Rのリサーチャーなのである。

次にA B Rの方法について。A B Rでは、美的 (aesthetic) ・芸術的な描写を用いる。だが、美的・芸術的といっても、写真や音楽やダンスやそれらの組み合わせといったような狭い範囲の芸術的手段を用いて評価するという意味ではない。A B Rの目的を簡潔に述べると、「言葉では言い表せない物事の意味を表現すること」⁽⁴⁶⁾

だとされるが、そのためには、A B Rの方法は、このような狭い範囲の芸術的手段によるのではなく、美的な考察を起源とする描写法にこそ、A B Rの方法が接近する可能性が探られねばならない。A B Rは一種の発見的方法 (a heuristic) といえる⁽⁴⁷⁾。発見的方法では、いつも正しく対象をとらえることは必ずしも保証されないが、ある程度のレベルで対象を正しくとらえることができるというのが特徴である。また、一般的に、発見的方法は対象をとらえるまでの時間が少なく済むことも特徴である。

第三に、A B Rのリサーチの役割について。「演劇の分野で用いられているストラテジーやテクニックがA B Rにおいて豊富に採用されうる」⁽⁴⁸⁾というように、演劇とA B Rの共通点から、リサーチの役割は演劇とのアナロジーで語られる。

演劇をよく理解するには、演劇そのものを見るだけでは不十分である。例えば、テーマ性やアイデアをもった複雑な演劇の場合には、たった一度おおまかに見ただけでは十分に正しく理解できないこともしばしばである。そこで、演劇では、観客に理解を促すために、まず事前の説明としてパンフレットやプログラムなどを用意して、観客にそれらを事前に読ませることで演劇を理解するのに必要な知識を知らせる。さらに、上演の後には討論の時間 (discussion segment) が設けられ、討論のリーダーの主導の下、その場にいる人たちと、この演劇のメッセージは何だったのか、説得力があったかどうか、自身や身近な人々の生活に関係したかなどを話し合うことを通して演劇に対する観客の理解を促す。

A B Rにおいても、演劇を見る場合と同様である。つまり、リサーチしようとする対象のパフォーマンスそのもののみを見るだけでは対象について十分に理解できないことがしばしば起こり得ると考えられる。そこで、対象をよりよく理解するために、A B Rのリサーチは、対象のリサーチにおいて、「前置き (preamble)」、「上演 (play)」、「話し合い (conversation)」の三つのプログラムを意識的に行う必要がある⁽⁴⁹⁾。ここにおいて、A B Rのリサーチの役割はもはや観察や解釈といった一方的な調査を実行するという役割から、捉えようとする対象や周りのオーディエンスをも巻き込んだ双方向的、多方向的な話し合いを組織するという役割へとシフトしている。これは、ポスト実証主義的アプローチにもとづくパフォーマンス評価の方法に一つの示唆を与えている。

6. おわりに

現在、わが国の教育評価においては、パフォーマンス評価への期待が高まっている。ペーパーテスト以外も含

めた広い意味でのパフォーマンスを評価の対象とするパフォーマンス評価だが、実証主義的アプローチと結びつけて議論されることが多かった。実証主義的アプローチの問題性は、わが国では「羅生門的アプローチ」およびアイズナーやスクリヴァンの批判が主導して、今や広く認識されている。この認識のもと、パフォーマンス評価で評価基準表を用いる場合であっても、途中で評価基準の変更を可能とするなど、実証主義的アプローチの問題性に対する自覚の下、実証主義的アプローチとポスト実証主義的アプローチとの統一の様相を描く方策が模索されている。

だが、実証主義的アプローチにポスト実証主義的な視点を取り入れるという対応の仕方には限界があり、純粹にポスト実証主義的アプローチの立場からパフォーマンス評価の方法を検討することが求められている。ポスト実証主義的アプローチは実証主義的アプローチと比べて、採点の一貫性が確保しにくいといった点で、普及に時間を要しているが、A B Rで示されたように、以前よりも明確な形で評価の方法や評価者の役割などが明らかにされるなど、その可能性は高まっている。今後、ポスト実証主義的アプローチからパフォーマンス評価の方法を検討する意義は大きいといえよう。

注

- (1) 松下佳代 (2007) 『パフォーマンス評価—子どもの思考と表現を評価する—』日本標準、7頁。
- (2) 同上。
- (3) 実証主義とは、理論や仮説や命題を経験的事実によって検証しようとする立場である。実証主義においては、超越的なものの存在 (経験的事実によって根拠づけられないもの) は認められない。すなわち、実証主義の評価とは、仮説としての評価基準 (基準) を用いて到達度を検証していくという評価の方法一般のことをいう。
- (4) ダイアン・ハート著、田中耕治監訳 (2012) 『パフォーマンス評価入門—「真正の評価」論からの提案—』ミネルヴァ書房、19頁。
- (5) 同上書、22頁。
- (6) 同上参照。
- (7) 同上書、23頁参照。
- (8) 同上書、24頁参照。
- (9) 同上書、25頁参照。
- (10) 1980年代のイギリスで開発された「プライマリー・ランゲージ・レコード (Primary Language Record)」と呼ばれる記録システムが、今やニューヨーク州、ニュージャージー州、カリフォルニア州の学校で試用されており、評価ツールとして一定の成功を収めている (ハート2012、pp.26-29参照)。
- (11) ダイアン・ハート2012、前掲書、p.32。
- (12) 同上参照。
- (13) 同上書、34-35頁参照。
- (14) 同上書、39-52頁参照。
- (15) 同上書、54頁。
- (16) 同上書、57頁参照。
- (17) 基本的に生徒が個人で取り組む課題で、通常、完了まで

- に数分しかかからない。
- (18) 生徒を巻き込んでチームやグループで作業する課題で、完成までにある程度期間を要する。
- (19) 長期にわたる多目的プロジェクト。長期プロジェクトでは活動や行程がカリキュラムの中に位置付けられている。
- (20) ダイアン・ハート2012、前掲書、69頁。
- (21) 同上書、78頁。
- (22) 田中耕治 (2010)『新しい「評価のあり方」を拓く―「目標に準拠した評価」のこれまでとこれから―』日本標準、23頁。
- (23) 天野正輝は相対評価の問題性を次のような点に要約している (天野正輝 (1993)『教育評価史研究―教育実践における評価論の系譜―』東信堂、pp.294-295)。1、相対評価が依拠した正規分布の原理に無理がある。2、相対評価は、教師に対して教育課程や授業の改善のための情報を提供しないし、子供に対しても学習への目安と励ましを与えない。3、相対評価は、児童・生徒にどんな知識をどれだけ教えねばならないかについて客観的基準をもっていない。4、相対評価のもたらす激しい競争意識が排他的な個人主義を助長し、学級づくり、仲間づくりを困難にする。5、能力は自然的法則に従って発達するという考え方は、子どもの発達における教育的働きかけの積極的役割を見落としてしまい、教師の側の日常的教材研究や指導方法改善への自覚と努力を失わせる。
- (24) 1975年前後から東京や京都において、到達目標・到達度評価についての理論的研究と実践が盛んに展開され (到達度評価運動)、今日に至るまで精力的に進められてきた。
- (25) 相対評価には、教師側の進路指導や生徒側の進路決定においてのニーズがあり、学期末の通知表などにおいては引き続き用いられた。
- (26) 2001年4月27日文科科学省通知の指導要録では、相対評価がなくなり、「目標に準拠した評価」に全面的に移行した。
- (27) ただし、現在のわが国における目標準拠評価をめぐるのは、そのオルタナティブ・アプローチとの統一の様相を描くことが模索されており、評価自体に評価基準づくりというプロセスを含めて途中での評価基準変更の可能性を拓くなど、厳密な実証主義に限定されない方法的柔軟性をもって議論がなされていることにも注目しなければならない。
- (28) 田中2010、前掲書、5頁。
- (29) 同上書、11頁。
- (30) 同上参照。
- (31) 同上書、5頁。
- (32) The Joint Committee on Standards for Educational Evaluation James R. Sanders.Chair (1994) *The Program Evaluation Standards 2nd edition: How to Assess Evaluation of Educational Programs*. SAGE Publications Inc., p.5-6
- (33) Ibid., p.5
- (34) Ibid., p.6
- (35) Ibid.
- (36) Ibid.
- (37) アメリカでは、スプートニク・ショック以後に莫大な費用を使って実施された多様な教育プログラムの評価を行う必要性にせまられた。そこでは即時に評価することが求められたが、1930年代のタイラー以後用いられていた実証主義的アプローチではその期待に応えられなかった。そこで、実証主義に依らないアプローチ (ポスト実証主義的アプローチ) が1960年代以降に多様に提唱され、目標準拠評価のオルタナティブ評価となった。
- (38) 田中2010、前掲書、24-25頁。
- (39) 同上書、25頁。
- (40) 同上書、25-26頁。
- (41) 同上書、26頁。
- (42) 田中は、自らが整理した (2) から (5) の批判について、これらを目標準拠評価との対立 (相容れないもの) として捉えるのではなく、目標準拠評価とそれらの批判から提唱されるオルタナティブとの統一の様相を具体的に描くことが必要であると述べている。その統一の構想に重要な示唆を与えるものとして、「目標準拠評価と個人内評価の内在的結合」や「真正の評価」論などに言及している (田中2010、pp.70-71参照)。
- (43) 近年の評価研究において、「利害 (stakes)」という概念が注目されている。これは評価の利害関係者の検討や評価結果のもたらす社会的影響の検討に評価研究の道を拓く。
- (44) Cf., Barone.T & Eisner,E.W. (2012) *Arts Based Research*. SAGE Publications Inc., pp.2-3
- (45) Ibid., p.43
- (46) Ibid., p.1
- (47) Cf., ibid., p.3
- (48) Ibid., p.59
- (49) Cf., ibid., 60