

# 教育評価の鳥瞰図（2）－アクティブ・ラーニングの評価研究

\*平 真木 夫

The Bird's Diew of Educational Evaluation (2) – A Study of Evaluation on Active Learnings

TAIRA Makio

## Abstract

The author first considered relation between backward design and curriculum management. Moreover, the author developed bird's view model of educational evaluation and based on that model pointed out failures of relative evaluation. The author secondary noted transition to active learning activities that becomes the focus of the next course of study in terms of educational evaluations, then introduced One-Paper Portfolio Assessment which is useful to evaluate active learnings.

**Key words** : ルーブリック

一枚ポートフォリオ

カリキュラム・マネジメント

## 1, 指導と評価の一体化とは

教育評価の話をする前に、そもそもなぜ評価が必要なのかその理由を考えてみたい。

例えば、授業を考える方法論として 逆向き設計理論<sup>1</sup> (Wiggins & McTighe, 2007) という概念が存在する。そこでは、始めに求められている目標を明確にし、修了時の結果の検証方法（評価方法）を決定し、学習活動と指導を明らかにすべきと主張する。つまり、指導の前に（1）目標を明確にする、（2）目標と照らし合わせて評価する、（3）指導を改善する、そして（4）すべての子どもに学力を保障するという方法論である。これらをまとめると、評価の観点と指導の目標とは一致しているべきで、学習の指導もすべての子

どもたちが目標に到達できるように行うべきであるということとなる。このような考え方はカリキュラム・マネジメントと関連していると考えられるが、究極的には、教育評価というものは、教育を改善するために行われる行為と定義することもできるであろう（西岡, 2008）。

このような評価のサイクルを行うために評価を行うタイミングも幾つかに分けられ、それぞれのタイミングで行う評価の性質も異なっている。例えば、学年の初めや単元の初めに行う評価として診断的評価（diagnostic evaluation）というものがあるが、そこではその後の学習に必要とされる知識・技能がどこまで確かであるか確認される。別の言葉で言い換えると、新しい単元や領域を学習するために、どの程度準備が

---

\* 教職大学院

- 1 ウィギンズ (Grant Wiggins) とマクタイ (Jay McTighe) が提唱し始めた理論である。思考力・判断力・表現力といった高次の学力を子どもたちに身につけさせるため、カリキュラムを評価から「逆向き」に設計することを提唱している。
- 2 学習に関して準備ができていないかどうか、その準備性を意味している。例えば、折れ線グラフを学習する単元の場合には、表の読み取りができていなければならない。他にも身体的な成熟による準備性を意味することもある。

できているか、レディネス (readiness)<sup>2</sup>などを診断するために、テストや調査が必要とされると言えるであろう。

次のタイミングは授業中に実施される形成的評価 (formative evaluation) がある。形成的評価は必ずしも授業時のあいだに行われるものに限定されないが、基本的には授業の途中のそこまでの成果を把握し、その後の学習を促すために行う評価と定義できる。つまり、学習者が現段階でどの程度教育目標を達成できているかを見るためのものと換言できる。具体的には、授業の最後に実施される小テストなどが形成的評価の例として挙げられる。このような形成的評価を行うことにより、学習者は自分自身の到達度を知り、学習活動を調整し、的確な復習を行うことができるようになる。逆に、教える側にとって考えると、到達度の低い学習者への個別対応が可能になる他に、カリキュラムや指導方法、教材などを改善できるという利点がある。

そして、評価が行われる最後のタイミングとして、

単元終了後、学期の中間、学期末などがあるが、そこで行われる評価が総括的評価 (summative evaluation) である。総括的評価は単元終了後など、最後に学習の成果を総合的・全体的に把握するために行う評価であり、指導内容全体を対象とするテストなどを行い、学習者の最終的な到達度を確認することになる。つまり、評価を行うタイミングを最初、中間、終盤と考えると、それぞれ診断的評価、形成的評価、総括的評価に分けられるということである。

## 2. 教育評価の鳥瞰図

教育評価において重要な評価として、絶対評価 (absolute/achievement-based evaluation) と相対評価 (relative evaluation)、そして自己評価 (individual evaluation) が挙げられる。それぞれの関係を、数量化のしやすさの次元と評価基準の外在性の次元の2つに基づいて図1のように整理可能である。

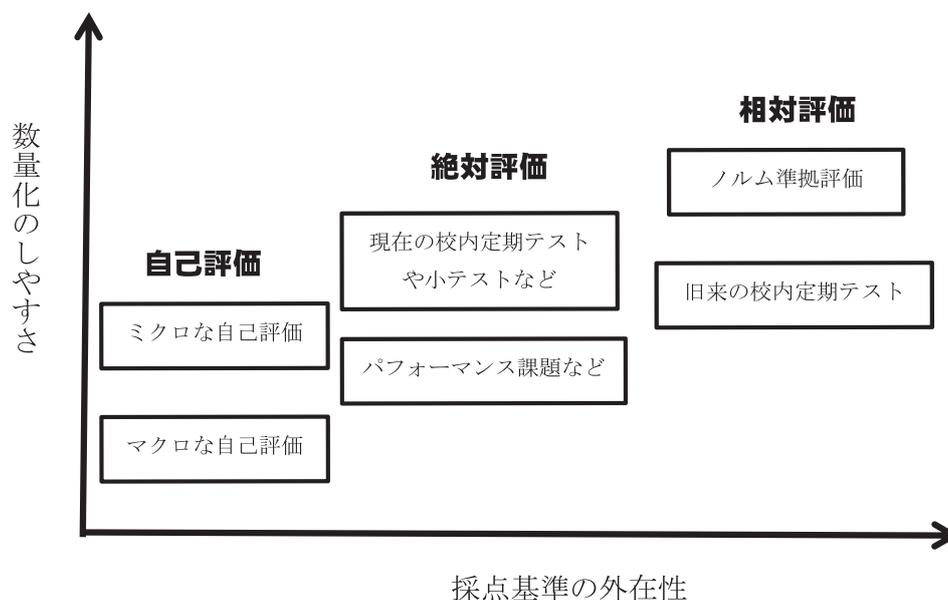


図1 教育評価の鳥瞰図

この図は平 (2002) の「図1 各種評価技法の見取り図」を改良したものである。

基準の外在性とは、評価基準 (学習目標、到達目標など) が学習者の中にあるのか、外にあるのかを意味している。例えば相対評価は準拠集団の中の位置に基づいて評価されるため、極めて外在性が高いと言えるだろう。それに対して自己評価では評価基準が自分の中にあるため、外在性は極めて低いと言えるだろう。

絶対評価 (到達度評価) の評価基準は、ある種の学問体系に照らし合わせて作成されている。その意味では評価基準は学習者の外にあるが、最終的には学習者の内部に取り込まれることが期待されており、中間的な位置に位置づけられると言えるだろう。

数量化のしやすさとは、統計学的な用語で表現す

ると、得点の空間的位置情報の多さ、厳密さを意味している。例えば、ノルム準拠評価で用いられる問題では、統計学的な標準化が施されており、予備校の模擬試験などのように偏差値といった数量的な情報が豊富に得られる。それに対して、絶対評価におけるパフォーマンス課題では後述するルーブリック(rubric: 評価指標)のような言語的に表現された行動目標や典型例(アンカー)とのマッチングによって評価されるため、数量的な情報を得ることは難しい。更に後述するマクロな自己評価のようなキャリア教育<sup>3</sup>における自己評価では、自己の進歩や変化を評価することになり、定量的変化というより定性的な変化と表現する方が適切になってしまう場合もある。

## 2. 1 相対評価

図1に示したように相対評価の特徴は特定の準拠集団(ノルム)を設定し、その中での位置に基づいて成績が評価されることにある。そのため、準拠集団を設定できれば、あとはその位置情報だけで自動的に成績がつけられる、ある種「科学的な」成績評価システムと言える。しかしながら、相対評価に関しては様々な批判や欠点が指摘されている。

第1に、旧来の校内定期テストでは準拠集団の人数がせいぜい1クラス30～40名程度であり、予備校などが主催している模擬試験のように千人といった大規模な場合とは想定される分布が異なっている。例えば、一般的に通常のクラスでは学力得点は双峰性の分布となることが多いし、同じ学年でもクラスごとで得点の分布が異なっていることもある。そのため他のクラスで成績が5段階評価で3であった生徒が別のクラスでも確実に3になる保証はない。それに対して、予備校が主催している模擬試験では様々な学力層の生徒が受験するため、得点が比較的きれいな正規分布を形成していて、正しく相対評価が可能になっている。つ

まり、旧来の校内定期テストのように、少ない生徒数を準拠集団とした相対評価は理論上問題があったと言えるだろう。

第2に、相対評価では到達度と関係なく成績をつけることが可能であり、学力が保証されているか定かではないという批判も挙げられる。これは学校側が保護者や社会に対して学力保証の説明責任を果たしているか分からないということでもある。また、順位を過剰に重視することにより、試験にあわせて学ぶようになるという欠点も生じる。極端な例を挙げると、センター試験の数学である領域を捨て、他の領域の学習に集中するという学習方略はよく見られる。このような問題があるため、平成13年度の指導要録<sup>4</sup>の改定から、学力保証を果たしていることを示すために相対評価から絶対評価へと移行したのである。

第3に、相対評価と関連した問題として、順位を上げることが学習目標となってしまうと様々な問題が生じる可能性が指摘されている。例えば、Turnerらの調査によると(Turner, et.al, 2002)、競争を意識した学習目標構造(performance goal structure)のクラスでは学びからの逃避(self-handicapping)が見られたという。それに対して、達成度を目標としたクラス(mastery goal structure)では学びからの逃避はあまり見られなかったことが報告されている。

なお、評価と関連した心理的現象として最も重要なものとして、テスト不安(test anxiety)を挙げることができるであろう。テスト不安とは一言で言うと、「自分が他人よりも良くないと評価されてしまうことは、不快な経験であり、自尊感情が損なわれ、試験が行われるたびに情緒的な緊張が起こる現象」である。例えば、高不安の生徒は、テストと無関係な反応をして、成績が悪化してしまうことも知られている(セルフ・ハンディキャッピングなど学びからの逃走が相当している)(Stigginns, 2001)。その一方で、低不安の生徒は、評価場面で挑戦することによって、成績が一層良

3 文部科学省の答申「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」(平成23年1月31日)によると、『人が、生涯の中で様々な役割を果たす過程で、自らの役割の価値や自分と役割との関係を見いだしていく連なりや積み重ねが、「キャリア」の意味するところである。…中略…このような、一人一人の社会的・職業的自立に向け、必要な基盤となる能力や態度を育てることを通して、キャリア発達を促す教育が「キャリア教育」である。それは、特定の活動や指導方法に限定されるものではなく、様々な教育活動を通して実践される。』とキャリア教育が定義されている。

4 指導要録は学籍の記録と共に指導の記録が書かれている公文書であり、文部科学省が定めた書式によって作成される。学校独自で作成される通信簿・通知表とは異なって、高校入試や大学受験のときに提出される内申書に記載する事項にほぼ相当している。詳細は文部科学省のホームページに例示されている。

くなることも知られている。つまり、テストを完全にやり遂げることによって不安を解消しようとするということである。

## 2. 2 絶対評価

先述したとおり、指導要録の改定にともなって学校現場では絶対評価によって成績がつけているが、このとき基本となる概念は、到達目標の達成度をもとにした評価ということである。したがって、現行の絶対評価は到達度評価または目標に準拠した評価とよばれるが、どの様に到達目標を設定し評価するかという問題が生じることになる。本節ではこのことについて考察していきたい。

絶対評価の問題を論じるときに必ず出てくる問題が「キジュン」という言葉の使い分けである。例えば、算数の単元目標として「異分母の分数の足し算ができる」といった目標設定が可能であるが、このような達成キジュンは規準(通称ノリジュン)とよばれる。そして、この目標を具現化した問題群を「どの程度」できるようになったら十分満足できたと判断するかという量的なキジュンは基準(通称モトジュン)とよばれる。

しかし、到達目標自体が量的に明確に厳密に測ることができないような場合もある。例えば、小学校算数で「コンパスと定規を用いて正三角形を作図する」といった課題を設定することは一般的であろう。他にも修学旅行に行ったときに訪れた神社についてレポートを書かせる課題も一般的と言えるであろう。これらの課題も到達度に基づいて評価されるべきであるが、100点満点で厳密に評価することは難しいはずである。このような課題は総称してパフォーマンス課題とよばれるが、課題を評価するときに活用される評価ツールとして先述したルーブリックが知られている。

例えば、先ほどの算数の正三角形の作図問題をABCの3段階で評定する場合を例にとってルーブリック作成を説明しよう。B評価として「コンパスや定規の使い方、作図の手順に多少の多少のもたつきがあるが作図できる。辺の長さにも多少のずれがある」といった具合に言語的に水準を記述し、そして、可能であればB評価のアンカー(ベンチマーク: benchmark)としてその典型例も添付する。この様にルーブリックを活用した絶対評価、目標に準拠した評価では、指導の前

に目標を明確にし、目標と照らし合わせて評価できるという利点がある。

ただし、全ての課題や単元にルーブリックを設定することは現実的ではない。実際にレポート課題やポスターなどアンカーとなる典型例を教員一人で選定することは困難で、可能であれば複数の教員で協同でルーブリックを構築すべきである。実際には困難かもしれないが、アンカーと呼ばれる典型例を選定する段階から複数の教員で協力しながら教材研究を行い、その中でルーブリックを作成していくべきであろう。

なお、絶対評価に関係する興味深い現象として、評価のインフレ現象を指摘することができる。これは実際に相対評価から絶対評価へと評価方法が移行したことによって、横浜市の教育委員会が評価のインフレ現象が生じたことが指摘されている(毎日新聞2004年5月25日)。例えば、ある中学校では全9教科で5の平均が4割を超え、評定平均は2年生で3.79であった。これと同種の現象はアメリカではレイク・オビゴン効果(Lake Wobegon effect)として知られている(Phillips, 1990)。全米50州の実力テストの結果を行ったところ、ほとんど全ての州で業者が設定した標準平均得点を上回っていたという現象である(どの州も平均より高いという結果)。これらは、評価基準の設定が各学校(各州)に委ねられていたため生じた現象で、到達目標と評価のモデレーションの必要性が示唆されている。

## 2. 3 自己評価

自己評価は個人が独力で自らの変化を評価することと考えられてきたが、子供たちはそこまで強力な自己を初めから獲得しているのであろうか。例えば、自分の行動や状態を自分で客観視する行為はメタ認知とよばれるが、初めから一人でその様なメタ認知ができるのであろうか。この問いかけに対する答えは、レイヴ(Jean Lave)とウェンガー(Etienne Wenger)が提唱する正統的周辺参加という概念で答えることができるであろう(Lave & Wenger, 1993)。

日本では体育会系の部活動などがその典型例だと思われるが、部活を行った最後の時間にミーティングが行われることが多い。そこでは顧問の教員やキャプテンが各部員のプレーにコメントすることがある。つ

まり、個々のメンバーのメタ認知的判断は他者がやってもかまわないということであり、その領域の初心者にとって他者の判断基準の取り込みを通じて自らのメタ認知能力を育成することが望まれるということである。この様な考えを一般の教科学習に当てはめて考えると、ペア学習などを行い、それぞれの学びを観察することにつながる。その結果、自分の弱点などを把握して予習・復習を促したりする。この様なメタ認知を中心とした自己評価はマイクロな自己評価と表現可能で、勉強する範囲を考えるためにもある程度の正確さが必要とされる。

これに対して、キャリア教育の場面などで行う自己評価は、たとえ他者の手助けがあったとしても厳密にはできないであろう。例えば、職場体験や一般の教科学習などをとおして自分の適性を考え、将来的な社会的自立を考えさせるのがキャリア教育の一般的な形式である。このときに行われる自己評価は上記のマイクロな自己評価とは異なってマクロな視点で行われるものであり、厳密には数量化できないはずである。例えば、小学校教員を目指している中学生が国語や体育を勉強

していたとして、それがどの程度あとで役に立つか、中学生の時点では予測もつかないだろう。

以上の考察から分かることは、自己評価にはマイクロなものと同様にマクロなものがあり、図1に示したように、それぞれ数量化のしやすさに違いがあるということである。

### 3. アクティブ・ラーニングに基づいた学習の評価

アクティブ・ラーニング(以下ALと省略)の定義は研究者によって多様で、一様に定まらないことが逆にALの特徴となっているが、「一方的な知識伝達型講義を聴くという(受動的)学習を乗り越える意味での、あらゆる能動的な学習のこと。能動的な学習には、書く・話す・発表するなどの活動への関与と、そこで生じる認知プロセスの外化を伴う」と定義可能であろう(溝上, 2012)。また、この様な視点に立つと具体的な学習活動としては山地(2011)が示したようになる(図2)。

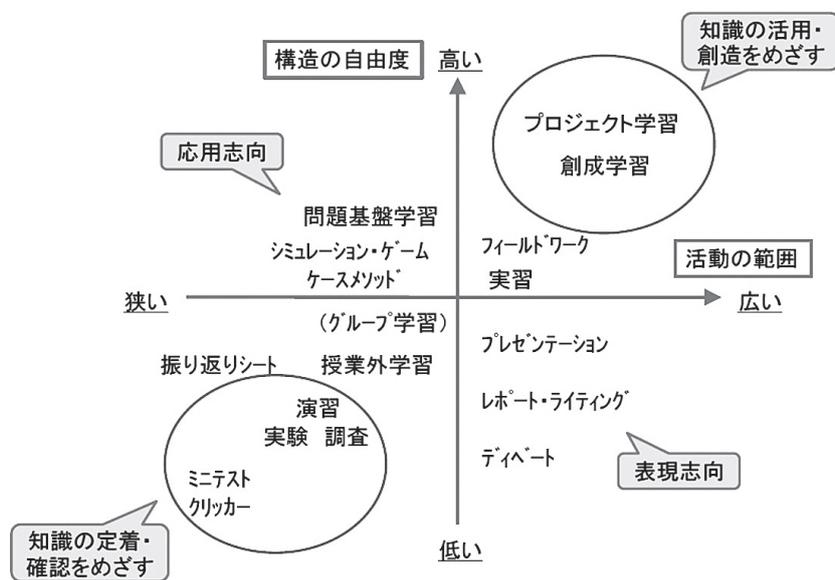


図2 アクティブ・ラーニングのタイプ分け

この図は山地弘起(2014)「アクティブ・ラーニングとはなにか」における『図1 アクティブ・ラーニングの多様な形態』からそのまま引用した。大学におけるALを念頭に作成されているが、基本的な考えは小・中・高でも適用可能と考えられる。  
([http://www.innov.nagasaki-u.ac.jp/teacher/files/Int\\_yamaji.pdf](http://www.innov.nagasaki-u.ac.jp/teacher/files/Int_yamaji.pdf))

5 クリッカーとは、講義中に端末のボタンを押すことによってその瞬間の理解度を測るための機器である。例えば、受講者は講義内容を聞きながら「理解できた」または「理解できなかった」といった反応を、クリッカーのボタンを押していつでも意思表示することが可能で、その結果をリアルタイムで表示可能になっている。

山地 (2011) が示した A L の例は大学教育を念頭に置いたものであるが、溝上 (2012) が定義した様に能動的な学習プロセスを包含したものになっている。例えば、クリッカー<sup>5</sup>のように扱っている道具は異なるが、基本的には小学校、中学校、高等学校でも類似の活動を取り得る。実際に酒井 (2013) から中学生を対象にクリッカーを活用した実践研究が報告されている。

図 2 に示されているように、クリッカーを活用した様な実践は能動的な活動をとおして知識の定着・確認を目指した A L と言えよう。それに対して、第 1 象限に位置するようなプロジェクト学習のような課題探究型の学習活動は、知識の活用をとおして創造性を目指すものであり、構造の自由度が高く第 3 象限に分類されるような学習活動と比べて評価を行うことは困難であると考えられる。

実際に A L の一種である課題探求型の学習の評価について現職教員を対象とした講習会で質問したところ、探求活動自体は指導できるもののそれを評価することは困難であるという回答が得られた [ $F(1, 64) = 22.01, p < .01, MSE = .443$ ]。指導と評価の一体化という文脈で考えると、結局、A L の評価の仕方が分からないということは A L で何を行うのが分からないということを実際には意味しており、深刻な状況にあることが示唆される<sup>6</sup>。

課題探求型の学習はパフォーマンス課題を活用した学習の 1 種と考えられるが、パフォーマンス課題の評価にはポートフォリオ評価<sup>7</sup>の実施が要求される。例えば、既存の評価はペーパーテストに偏重しているが、そこでは評価の方法とタイミングを固定して、そこから捉えられるもののみを評価してきた (石井, 2015)。しかし、パフォーマンス課題を用いた評価では、課題を遂行するために必要とされる個別の事実に知識も評価されると同時に、実験の計画・立案に関

わる方法知も評価される (西岡ほか, 2013)。そこでは、理解の程度や、能力の熟達化の程度で評価されることになる。つまり、作品集であるポートフォリオを、これまでも繰り返しできたルーブリックを作成し質的に評価することになるのである。

しかしながら、ポートフォリオ評価自体には幾つか問題が指摘されている (堀, 2013)。

1. 活用できないような雑多な情報が膨大に含まれている。多種多様な情報を集めることは必ずしもよいことではない。
2. 多様な情報 (作品等) を取捨してポートフォリオとして評価することが望まれているが、情報の選択の適否を誰がいつやるかが不透明である。
3. 雑然としているので、学習の前・中・後の評価を取り扱うのが難しい。
4. 自己評価における学習目標が曖昧で、自己評価を行うことの必然性が伝わりにくい。
5. 学修による変容、特に自己の変容をとらえにくい。
6. 生徒の問題点や分からなかったことなどが見えにくいいため、授業や学習に生かすにくい。

このような欠点を解決するために、堀哲夫が開発した一枚ポートフォリオがある (堀, 2013)。

図 3 は堀 (2013) を参考にして平が作成したものである。①の診断的評価に相当する問いと②の総括的評価に相当する問いは原則的に類似のものになっていることが望まれる。例えば、研究課題名として「地球に優しい発電」というテーマがあったとき、①の診断的評価となる問いは「地球に優しい発電の必要性について簡単に説明しなさい」といった問いが考えられるだろう。他にも、一般の教科学習を想定すると、「メダカの繁殖に必要な条件を考えなさい」といった問いも考えられるだろう<sup>8</sup>。①の診断的評価の場面での記述

6 A L の取り組みについては、校種の違いも調査の結果から明らかにされている。A L への取り組みそのものは、小学校、中学校では特に支障は無いと考えられているが、高等学校においては統計的に有意な差が示唆された [ $F(2, 64) = 3.33, p < .05, MSE = .655$ ]。

7 文部科学省の用語集の定義によると以下ようになる。「学生が、学修過程ならびに各種の学修成果 (例えば、学修目標・学修計画表とチェックシート、課題達成のために収集した資料や遂行状況、レポート、成績単位取得表など) を長期にわたって収集し、記録したもの。それらを必要に応じて系統的に選択し、学修過程を含めて到達度を評価し、次に取り組むべき課題をみつけてステップアップを図るといふ、学生自身の自己省察を可能とすることにより、自律的な学修をより深化させることを目的とする。従来の到達度評価では測定できない個人能力の質的評価を行うことが意図されているとともに、教員や大学が、組織としての教育の成果を評価する場合にも利用される。」

は生徒によって様々であるが、基本的にはその時点で持っている既有知識の確認となり、結果的にキーワードの羅列や素朴な意見の箇条書きとなることが多い。

それに対して、②の総括的評価の場面では、科学的なアイデアや用語を駆使した理論的な記述となる(ことが望まれる)。

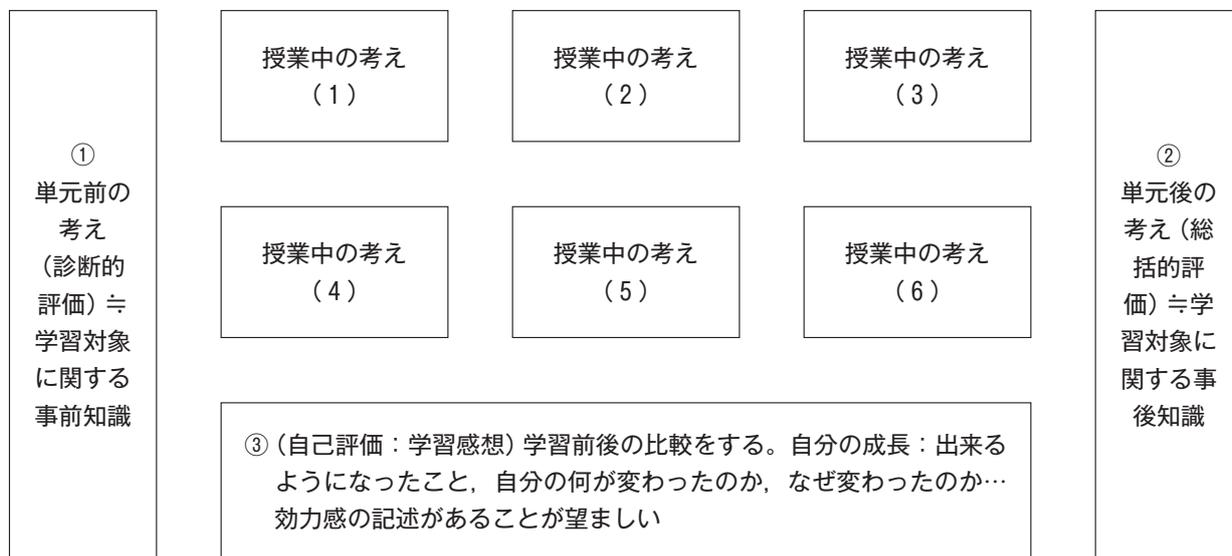


図3 一枚ポートフォリオのサンプル

この図は堀哲夫(2013)「教育評価の本質を問う 一枚ポートフォリオ評価 OPPA」p23を参考に平が作成した。

そして、③の学習後の自己評価では、①と②の答えを比較して自分の中で何が変わったのか、どこが成長したのかメタ認知を行うことになる。これは図1における用語で考えるとマイクロな自己評価に相当するが、一枚ポートフォリオの重要な特徴は各授業の中で何を学んだか振り返るところにある(「授業中の考え1～6」)。このことによって、その授業の振り返りというメタ認知を促進させると同時に、それ以外に教師としては生徒の誤解をいち早く察知できるようになり、次の時間の授業で修正ができるようになるという利点もある。

一枚ポートフォリオの利点は以下のようにまとめることができるだろう。

1. 自分自身が評価することによって、自分の成長や効力感を得ることができる。
2. その時間での重要なことを書かせることによって、

次に学習する目標を持たせることができる。

3. 思考や認知過程を内省し、それを書かせることによって外化できる。
4. 内省させることによって、自己評価やメタ認知能力を向上できる。
5. 外化(文章として明示化)させることによって、思考の情報処理の負担が軽減できる(結果的に高次の判断であるメタ認知をする余力ができる)。
6. 各生徒の記述を確認することで、授業の修正ができる。

しかしながら、一枚ポートフォリオを活用するにあたって幾つか気をつけなければいけないことがある。

例えば、一枚ポートフォリオを活用するにあたって一番難しいポイントが最初の診断的評価に相当する発問の作成である。この発問がその単元をまとめる根幹となる中心発問に相当するものであり、この発問が

8 ①と②の答え方として言葉で記述させる方法以外に、概念マップなど図を利用することも可能である。例えば、小学校5年生の理科「魚の誕生」の単元では、①の診断的評価の段階でメダカの雄と雌の関係までは描けても、その後の発育のプロセスまでは描けないはずである。最後の③の総括的評価の場面で概念マップを描かせた場合には、ミジンコなど小さな生き物を食べる部分も含めて描けるようになるはずである(三浦, 2016)。

考えられなければ一枚ポートフォリオそのものが成立しなくなってしまう。また、生徒たちに授業を振り返ることができるだけの学力、文章力が必要とされることも重要な要素であろう。実際に、少なくとも小学3年生以上でないと文章を用いて振り返るような行動は難しいと思われるし、何を振りかえさせるか指示の工夫も必要とされる。そして、一枚ポートフォリオを書かせた場合の返し方の時間的コストの問題もある。全員のシートに細かく指導の言葉を書く必要は無く、認め印を押して二重丸やアンダーラインを引くことでチェックしたことを表すことは可能であるが(重要な間違いなどは個別にコメント意見を書くべきだが)、毎時間全ての生徒にその様な対応が可能かは担当している生徒数に依存するはずである。

以上のような注意点があるものの、利点として挙げたように思考を外化するための優れたツールであるのは確かであり、ALを促進するために役立つと考えられる。

#### 4. カリキュラム・マネジメントと教育評価

本稿の始めて評価とカリキュラム・マネジメントとの関連性について軽く触れたが、文部科学省の答申「次期学習指導要領等に向けたこれまでの審議のまとめについて」(平成28年8月26日)によると、カリキュラム・マネジメントのあり方として以下のような特徴を示唆している。

- ①各教科等の教育内容を相互の関係で捉え、学校教育目標を踏まえた教科等横断的な視点で、その目標の達成に必要な教育の内容を組織的に配列していくこと。
- ②教育内容の質の向上に向けて、子供たちの姿や地域の現状等に関する調査や各種データ等に基づき、教育課程を編成し、実施し、評価して改善を図る一連のPDCAサイクルを確立すること。
- ③教育内容と、教育活動に必要な人的・物的資源等を、地域等の外部の資源も含めて活用しながら効果的に組み合わせること。

Wiggins & McTighe (2007) による逆向き設計理論からも、堀 (2013) による一枚ポートフォリオの活用

からも、答申の②に書かれているようにPDCAサイクルの一環として教育評価が位置づけられている理由が分かる。また、既に使い古されたスローガンとなってしまっているが、「指導と評価の一体化」という概念においても同様のPDCAサイクルが見て取れる。これらは広義のカリキュラム・マネジメントと言って良いだろう。

しかしながら、平成13年度に指導要録の改訂にともなって、各学校で評価規準表の作成が促され、各教育委員会にその規準表が提出されたが、現在、その規準表が現場で実際に活用されているとは思えない状況にある。その理由は幾つかあるはずだが、最も重要な要因は時間的な制約から一気に作成されたからではないだろうか。その結果、適切な教材研究から裏付けられた内容を構築できずに適当に作成されてしまったのではないだろうか。つまり、評価規準表は授業を振りかえるシステムを校内に担保するためのカリキュラム・マネジメントの道具であったはずだが、教材研究抜きに急いで作成されたため、結果的に誰も使わない評価規準が作成されてしまったと推察される。今後は、教科横断的なカリキュラム編成が重要になるはずだが(田村, 2016)、評価規準表の作成で示されたような失敗を回避できる学校経営の工夫が必要とされると思われる。

#### 文献

- 堀哲夫 (2013). 教育評価の本質を問う 一枚ポートフォリオ評価 OPPA 東洋館出版社
- 石井英真 (2015). 今求められる学力と学びとは一コンピテンシー・ベースのカリキュラムの光と影 日本標準
- ジーン レイヴ・エティエンヌ ウェンガー (1993) 状況に埋め込まれた学習—正統的周辺参加 佐伯胖 (翻訳) 産業図書
- 三浦彰文 (2016) 理科における知識の構造化について 宮城教育大学教職大学院平成28年度2年次中間発表会資料
- 西岡加名恵 (2008) 「逆向き設計」で確かな学力を保障する 明治図書出版
- 西岡加名恵・川地亜弥子・北原琢也・石井 英真 (2013). 教職実践演習ワークブック—ポートフォリオで教師力アップ ミネルヴァ書房
- Phillips W. P. (1990) The Lake Wobegon Effect. *Educational Measurement: Issues and Practice*, Vol 9 (3), 3-14.
- 酒井徹雄 (2013) クリッカーの活用について 北海道立教育研究所附属理科教育センター研究紀要, 第25号, 54-57
- Stigginns, R.J. (2001) Where is our assessment future and how can we get there from here? Lissitz, R. W. & Schafer, W. D.

- (eds) *Assessment in Educational Reform: Both Means and Ends*, Chap. 2, 18-48. Allyn & Bacon
- 平真木夫 (2002) 教育評価の鳥瞰図 ～ 評価コストを軽減するための一つの試案 ～ 宮城教育大学紀要 Vol.37, 341-356.
- 田村知子 (2016) カリキュラム・マネジメントの目的と方法 教育と評価, vol.62, 10, No. 742, 9-14.
- Turner, J.C., Midgley C., Meyer D.K., Gheen M.Q., and Anderman E. (2002). The classroom environment and students' reports of avoidance strategies in mathematics: a multi-method study. *Journal of Educational Psychology*. (94) , 88-106
- 山地弘起 (2014) アクティブ・ラーニングとはなにか アクティブラーニング事例集1 [http://www.innov.nagasaki-u.ac.jp/teacher/files/Int\\_yamaji.pdf](http://www.innov.nagasaki-u.ac.jp/teacher/files/Int_yamaji.pdf) (アクセス日: 2016年9月29日)

(平成28年9月30日受理)