

# 形成的アセスメントの理論的展開

安 藤 輝 次\*

1960年代にクロンバック, L. J. が提唱した“形成的評価”は、ブルーム, B. によってすべての学習者に一定程度まで完全に習得させるための方法として普及したが、それは、知識や技能を教師主体で教えるための指導法であった。

しかし、1989年、サドラー, D. R. が学習の目標や規準を学習者にも分かるように主張し、1998年に、ブラック, P. とウイリアム, D. は、様々な研究調査結果に関する効果量に着目して、教授・学習過程だけでなく学習者が自らの学びを評価し方向付ける“形成的アセスメント”を行えば、テストの成績がアップし、自尊心や動機付けにも好影響を及ぼすことを根拠づけた。それ以来、形成的アセスメントは、広く知られるようになり、最近では、より学びの機能を強調した“学習のためのアセスメント”としても展開され、また“中間的アセスメント”と同一視されるような混乱も生じている。

本稿は、このような形成的アセスメントの過程を辿りながら、理論的な整理を行い、①総括的アセスメントとの関連性、②フィードバックの手立て、③指導の枠組みにおける位置づけ、という点から現在の到達点を明らかにしようとするものである。

キーワード：形成的評価、形成的アセスメント、総括的アセスメント、  
中間的アセスメント、学びのためのアセスメント

## 1. 問題の所在

経済協力開発機構（OECD）の教育研究革新センター編著 Formative Assessment (2005) が有本昌広氏の監訳によって『形成的アセスメントと学力』（明石書店、2008）と邦訳されて以来、“形成的アセスメント”という言葉がわが国で知られるようになってきた。

Formative assessment と言えば、1970年代から80年代にかけて普及した Bloom, B. S. の完全習得学習 mastery learning で言う“形成的評価”を思い浮かべる人も多いであろう。

有本氏によれば、形成的アセスメントは、Bloom の理論の導入に大きな役割を果たしてきた梶田観一氏の著書『形成的評価のために』が「ある意味で先見の明に富んでいた」と言う<sup>1)</sup>。

しかし、この著書では、例えば、高次の能力について、「従来から用いられてきた小論文式のテスト、レポートの提出を求める方法、あるいはペーパーテストの一貫として、問題に自由記述の形で解答させるという方法が望ましい」<sup>2)</sup>と述べており、それと有本氏が解説する「学習プロセスに直接寄与できる『生徒自身が意思決定者である』という点が際立ってきた」<sup>3)</sup>とされる形成的アセスメントとは同一には捉え難いのではないだろうか。前者では、教師がテストを出題し、後者では、子どもが学びを主導するからである。

アメリカに目を転じてみると、形成的アセスメントに関して次のような誤解が生じている。1990年代のスタンダード運動と2002年施行の

\* あんどう てるつぐ 関西大学文学部初等教育学専修 教授

落ちこぼれをなくす法 (NCBL) とが結びついた結果、州レベルの大規模テストによって、子どもの達成度を測定し、低迷する学校では廃校勧告まで行なうなかで、大規模テストを受ける子どもの点数を上げるために中間的アセスメント interim assessment を導入しようとする学校も多く、教育テスト・サービス社 Educational Testing Service のように、大規模テストの問題例をデータベース化したものを形成的アセスメントと呼んでいる<sup>4)</sup>。また、ミシガン州教育局の担当者は、中間的アセスメントを指導前や単元終了時でも実施するという提案をしているが、これでは形成的アセスメントとの違いが鮮明ではない<sup>5)</sup>。

このように形成的アセスメントをめぐるのは、形成的評価と連続的に捉えたり、中間的アセスメントと混同したり、その定義について不明確なことが多い。本稿では、これらの点を明らかにした後、形成的アセスメントを学校においていかに展開するべきかという提案を検討したい。なお、引用する人名は、正確さを期するために英語で記し、人名の前に国を示さない場合には、アメリカ合衆国の研究者又は教育関係者等とした、また、形成的アセスメントは、幼稚園児から大学生まで適用されるので、学習者という言葉で統一した。

## 2. “指導と評価の一体化” から “評価と学びの連動” へ

1967年、Scriven, M. は、次のように形成的評価の必要性を訴えた。すなわち、評価に対する不安のために、カリキュラムの評価が目標に従属させられている現実には憂慮すべきであり、「厳しい評価をすることができる最大のものは、修正が望ましい課程の諸側面を確認することである」という Cronbach, L. J. の言葉を踏まえるならば、カリキュラム作成者は、今あるカリキュラムを実地に適用する中でその不十分さや欠陥を確認し、修正のためにフィードバックするという意味で形成的評価 formative evaluation を設けるべきであり、カリキュラムの最終的な評価を下す総括的評価 summative evaluation とは区別すべきであると<sup>6)</sup>。Scriven は、このようにカリ

キュラムを修正するために形成的評価を提案したが、その際に、Bloom, B.S. の教育目標の分類学にも注目をしていた<sup>7)</sup>。

そして、Bloom は、形成的評価の概念を「カリキュラムの作成のみではなく、教授活動や学習活動にも有効である」<sup>8)</sup> と考え、単元レベルで目標分析を行い、学びの途上で形成的テストを実施して、子どもの不十分な学びを確認し、矯正指導をするという完全習得学習を提唱し、広く知られるようになった。その後、Bloom は、小中学生2、3人を対象にチュータが完全習得学習の方法で教えると、子どもたちは、通常の授業の子どもたちとを比べて、達成テストの成績が偏差値 ( $\sigma$ ) 2 を上回っており、子ども同士の相互作用の教育効果も大きいというシグマ2問題を提起したが<sup>9)</sup>、Bloom の完全習得学習は、教師が何をいかに教えるのかを意思決定することを基盤としていた。

このように形成的評価は、Scriven でカリキュラムの開発・実施過程における情報を提供し、フィードバックすることに始まり、Bloom においてはすべての子どもを達成テストで一定程度まで完全に習得させる指導法として展開されたが、いずれも教師がカリキュラム作成や教授・学習過程における主体と位置づけられていた<sup>10)</sup>。つまり、形成的評価は、教師による“指導と評価の一体化”を企図していたのである。

ところが、1989年、オーストラリアの教育心理学者の Sadler, D. R. は、形成的評価 formative evaluation という言葉ではなく、子どもに“寄り添う”という意味で assessment という用語を使った論文「形成的アセスメントと指導システムのデザイン」を発表した。彼がそこで根拠にしたのは、情報マネジメントを専門とする Ramaprasad, A. が明らかにしたフィードバックの定義、つまり、フィードバックとは参照レベルとアクチュアルなレベルとの間のギャップを縮めるためのシステムであるという考え方である<sup>11)</sup>。

Sadler は、このフィードバックを形成的アセスメントの鍵であると捉え、参照レベルをスタンダードやカリキュラムの目標に見立てて、教師だけでなく子どもにも分かるようにし、アク

チュアルなレベルを複数の評価規準を使った子どものパフォーマンスの質的な判断に置き換え、両者のギャップを縮小するために、他の子どもが学習した事柄を評価し、矯正する活動に係わらせると、子どもの自己評価にも役立つと言う<sup>12)</sup>。そして、教師によるフィードバックだけでなく子どもにも自己モニタリングをさせることを形成的アセスメントと捉え、「多くの指導システムのゴールは、フィードバックから自己モニタリングへの移行を促進することである」<sup>13)</sup>と主張した。ただし、このような主張は、すでに幾つかの学校で実施済みとはいっても、「経験的にそれらの学校では、結果を生み出している」<sup>14)</sup>と述べるのみで、形成的アセスメントの明確な効果を示すことはできなかった。

そして、形成的アセスメントは、イギリスの Black, P. と Wiliam, D たちが 1998 年に著した 2 つの論文によって一躍脚光を浴びるようになる。彼らは、Sadler の所論に依拠しながら、「アセスメントと教室学習」と題する 70 頁弱の論文において、過去 9 年間の 160 以上の専門誌から約 580 の論文や図書の分担章を抽出し、そこから教室での評価が子どもの学習の改善に役立っているか否かという問題意識に立って約 250 編を抽出して、フィードバック、形成的アセスメントの方法・技術、システム、生徒の評価などのカテゴリーに分けて、先行研究の結果を検討・集約した。とりわけ、中学理科の探究学習では、子どもの達成にどの程度の影響力があるのかということの効果量 effect size で数値化すると 0.7 という顕著な結果が出ており、3000 以上のフィードバック研究報告をレビューした結果、子どものパフォーマンスの上昇に及ぼす効果量が平均 0.4 であることを突き止めた<sup>15)</sup>。効果量 0.4 というのは、形成的アセスメントを適用された学級の子どもの平均の達成度が普通の授業を受けた子どもの上位 35% と同じであるから、その教育効果は大きい。そして、彼らは、学習課題に対する子どもの動機付けの意義を唱え、進歩を見据えた教師とその子どもとの対話の質が大切であると訴え、今後の課題として、低学力の子どもへの対処法や形成的アセスメントと総括的アセスメントとの区別の不明瞭性や教師と子

どもとの責任の分担などがあると指摘した<sup>16)</sup>。

Black と Wiliam のもう一つの論文「ブラックボックスの内部」は、教師、子ども、テスト圧力、保護者の不安などのインプットと中等教育修了資格試験（GCSE）などの大規模テストの結果や教師の満足・不満足などのアウトプットの間にある教室で教師がどのように教え、子どもがどのように学んでいるのかということがブラックボックスになっていると指摘した後、次のように形成的アセスメントを定義した<sup>17)</sup>。

「“アセスメント”という用語は、教師が、そして、子どもが自分自身を評価する際に取り組むすべての活動を指しており、そこでは、彼らが従事する教授と学習の活動を修正するためのフィードバックとしての情報提供を行う。証拠を使って、これらのニーズにそった指導を実際にすると、それが“形成的アセスメント”となる。」

そして、彼らは、テストが達成度を高めるとどんな証拠もないと断言し、「形成的アセスメントの改善は、スタンダードをアップさせるという証拠があるか？」という問いについては、前掲論文のメタ分析結果に基づいて、「イエス」と明言し、形成的アセスメントの実験学級と通常の学級の比較研究から、次のように形成的アセスメントを特徴づけた<sup>18)</sup>。(1)教える人と教えられる人の間のフィードバックを高める新しい教授様式が必要であり、教室での実践の意義深い変化がなされなければならない。(2)様々なアプローチの根底には、効果的な学習を生み出すこと、特に子どもを能動的に関与させなければならない。(3)アセスメントを形成的に機能させるには、教授と学習を調整するために結果を活用しなければならない。(4)アセスメントが子どもの動機づけと自尊心に影響を及ぼす仕方や子どもを自己評価に係わらせることによる効果に注目すべきである。

とりわけ、形成的アセスメントを(4)の動機づけや自尊心の面から捉えるのは、Sadler も言及しなかった点であり、Brookhart によれば、形成的アセスメント研究の一つの特徴とみなされるべきであると言う<sup>19)</sup>。

そして、「改善の余地があるという証拠がある

か？」という問いに対しては、テスト至上主義に陥っている現状を憂い、「形成的アセスメントの改善方法についての証拠があるか？」の問いについては、子どもに自己評価の仕方を教えることや「子どもとの対話では、思慮深く、反省的で、理解を喚起し、探るように焦点化し、すべての子どもに自分で考え、表現する機会を提供するようにすべきである」と提言した<sup>20)</sup>。

Black と Wiliam の 2004 年の論文「ブラックボックスの内部の運用」では、1999 年から 2 つの学区の 6 つの中等学校の教員 24 人で始め、その後、初等学校の教員も加わった 5 カ年に及ぶキング・メドウエイ・オックスフォード・形成的アセスメント・プロジェクト (KMOFAP) で、前述の 3 番目の問い、つまり、形成的アセスメントを改善する方法について実践的裏付けを示した。それを要約すると、第 1 に、形成的アセスメントをうまく活用すると、大規模テストでさえ学習者の得点がアップする、第 2 に、そのような変容をもたらす要因として、①発問と答えの待ち時間を長くとり、②挙手なしにランダムに指名する、③誤答を不快に感じない教室風土を醸成することが必要であり、第 3 に、評定を介したフィードバックについては、出来た事柄と改善を要する点にコメントを添えて、子どもがコメントに反応する機会が学習機会であると捉えられ、学習効果があり、第 4 に、子ども相互のピア評価を定期的に行い、そこで出てきた誤概念を次の学習に活用したり、赤緑黄の信号カードで「良くできた」「普通」「分からない」ということを伝えたり、子どもに採点用のルーブリックの使い方を教えることが効果的であり、第 5 に、子どもに自らの学んだ事柄を振り返らせ、改善努力を促して、総括的テストを学習過程の一部として捉えさせるということである<sup>21)</sup>。そして、これらの提言の根底には、「効果的な学習の鍵は、子どもたちが新しいより強力なアイデアで形成される知識の再構造化を助ける道を見出すことである」ということを KMOFAP で学んだと述べ、構造主義的な知識観を表明し、④子どもが思考を表現する、⑤フィードバックを学習改善として位置づける、⑥子どもたちに協働させ、互いのアイデアを聴き、尊重する、

⑦正答か否かに焦点化するより、子どもに理解したことを表現させ、話し合わせて授業に積極的に参加させる、ことを力説した<sup>22)</sup>。

また、Black と Wiliam たちは、このような形成的アセスメントを介した様々な変化をもたらす方法について詳述した著書『学習のための評価』を 2003 年に出版し、そこで教師が努力すべき仕事を従来の手法で一生懸命にするのではなく、仕事の配分の仕方を組み換えて賢く仕事をするのが肝要であると提言した<sup>23)</sup>。そのポイントは、子どもの学習についての情報を集めて「長期のカリキュラム改善のために使うことができて、それは、子どもの学習を助けることはならない。教師にとっては形成的であるが、子どもにとって形成的ではないからである。」という言葉を噛みしめなければならないと言う<sup>24)</sup>。

要するに、形成的アセスメントとは、教師が教える目標を見据えて、学習者のアクチュアルなレベルを踏まえて教えた後、子どもが分からない点を子ども自身と一緒に見極め、スタンダードなどの参照レベルに照らして両者のギャップを縮めるために、教師は指導改善に、子どもは自分自身の学びの向上に努力するという意味で、“評価と学びの連動”がなされているのである<sup>25)</sup>。欧米でこのような形成的アセスメントの考え方が多くの教育関係者の関心を集め、広く受け入れられたのは、Black たちによれば、初等中等学校の実践的な証拠に裏付けられており、教師としての職能開発の中でスタンダードに関連した大規模テストの成績アップということに高い優先順位を置いていたからであると言う<sup>26)</sup>。

以上、形成的評価から形成的アセスメントへの変遷、つまり、指導と評価の一体化から評価と学びの連動への移り変わりの過程を辿った。目標に近づけるために学習を形成するという意味で、いずれも「形成的 formative」という言葉を使っているが、形成的評価と言う評価は、evaluation の訳語であり、教師が子どもの学びを値踏みし、評定するという意味合いがあるが、形成的アセスメントは、formative assessment であって、assess の語源が“子どもに寄り添う”ということから分かるように、教師が子どもたちの学びを評価する際には教師と子どもとの単な

る対話ではなく相互作用をし、互いに影響を及ぼし合うことを含意しているということである。

### 3. 学習のためのアセスメント

Black たちは、教育政策を提言するアセスメント改革集団 Assessment Reform Group (ARG) に形成的アセスメントの必要性を訴え、ARG は、1999 年のパンフレットにおいて Black たちの言う形成的アセスメントを①学習者への効果的なフィードバック、②学習者の学習における積極的な関与、③評価結果を説明するための指導調整、④学習者の動機づけと自尊心への大きな影響、⑤学習者が自己評価し、学習改善法を理解する必要性、と概括した<sup>27)</sup>。そして、形成的アセスメントを「学習のためのアセスメント assessment for learning」と言い替えて、「学習に関するアセスメント assessment of learning」である総括的アセスメントと区別し、学習者の困難や誤謬を発見するという否定的意味がある診断的評価という用語は使わず、むしろ学習者の強みと弱みの両面に目を向け、次なる指導と学びを展望する意味合いで「学習のためのアセスメント」を使いたいと述べた<sup>28)</sup>。さらに、ARG の 2002 年のパンフレットでは、学習のためのアセスメントを「学習者と教師が今の学習者が学習している場所と次に進む必要のある場所とそこへの行く一番良い方法を決定する際に使うための証拠を追求し、解釈する過程である。」と定義した<sup>29)</sup>。そして、Black と Wiliam は、最近では、教師と学習者だけでなく学習者の仲間からも捉えて、表 1 のように、学習のためのアセスメントを 3 側面によって特徴づけている<sup>30)</sup>。

学習のためのアセスメントは、その後、イギ

リスでは ARG の Broadfoot, P.M. によって、アメリカでは教育コンサルタントの Stiggins, R.J. によって普及され<sup>31)</sup>、前述の定義に示すように、学習者が今いる場所（フィードアップ）、学習者が行く場所（フィードバック）、そこへの行き方（フィードフォワード）という 3 分類は、広く知られるようになった。そして、Stiggins は、「学習のためのアセスメント」を総括的評価である「学習に関するアセスメント」と表 2 のように対比させて、その区別を明確にした<sup>32)</sup>。

このように形成的アセスメントは、「頻繁なテスト」という誤解を避けるためにも<sup>33)</sup>、学習のためのアセスメントと言い替えられ、欧米各国で学校における実践的裏付けを伴いながら、理論的根拠を積み重ねて、様々な定義がなされてきた。例えば、Brookhart, S. M. は、教師より前に学習者を据えて、次のように学習者重視の定義をした<sup>34)</sup>。

「形成的アセスメントとは、学習者と教師が(1)学習目標への焦点化、(2)フォーマルまたはインフォーマルなアセスメントを使って学習目標に関して現在の学習物を収集し、(3)学習者が学習目標に近づくための措置を取る（例えば、教師が指導法を、学習者が学習法を調整する）ような場合に、教師と学習者が従事する進行中の過程である。」と。

他方、中間的アセスメントも重視する全米州教育長協議会 The Council of Chief State School Officers (CCSSO) とも関係が深い Popham, W. J. は、Brookhart ほど子どもを前面に出すこともなく、次のように形成的アセスメントを定義した<sup>35)</sup>。

表 1 形成的アセスメント（学習のためのアセスメント）の 3 側面

	学習者が行く場所	学習者が今いる場所	そこへの行き方
教 師	学習意図の解明と成功のための規準の共有	学習の証拠を引き出すような効果的な教室討論、活動、課題の組み合わせ	学習者を前へ進ませるフィードバックの提供
仲 間	学習意図と成功のための規準の理解と解明	互いの指導資源として学習者を活かすこと	
学習者	学習意図と成功のための規準の理解	自分自身の学習としての学習者を活かすこと	

表2 学習のためのアセスメントと学習に関するアセスメントの比較

	学習のためのアセスメント	学習に関するアセスメント
評価の理由	達成を高めて、学習者がより高いスタンダードを満たすことを助ける。進行中の学習者の成長を支援し、改善する。	個人や集団の達成やスタンダードの習得を文書化する。報告の目的で時間上のある時点で達成の現状を測定する。説明責任を果たす。
被評価者	学習者自身	他の学習者たち
評価の焦点	学習者たちがスタンダードに向けて形成するのを助ける教師が選んだ特定の達成のターゲット。	学校、教師、学習者たちが説明責任を果たすための達成のスタンダード。
評価する時	学習中の過程	学習後のイベント
主たる使用者	学習者、教師、保護者	政策立案者、プログラム計画者、指導主事、教師、学習者、保護者
典型的な使用法	学習者たちに達成を改善するための洞察を提供する。教師が学習者のニーズを診断し、反応するのを助ける。保護者が長期にわたって進歩を捉えるのを助ける。保護者が学習を支援するのを助ける。	学習者のコンピテンスを証明する。達成にしたがって学習者を選別する。進級を卒業の決定。評定。
教師の役割	スタンダードを教室の目標に変換する。目標について学習者に知らせる。評価をする。結果に基づく指導を調整する。学習者に文章や言葉によるフィードバックをする。評価に学習者を関与させる。	結果の正確性と比較可能性を確保するためにテストを注意深く実施する。学習者がスタンダードを満たすのを助けるために結果を利用する。保護者に結果を報告し、通知表のために評価をする。
学習者の役割	自己評価し、進歩を辿る。ゴール設定に貢献する。教室の評価結果にしたがって、次によりよくできるようにする。	スタンダードを満足するために研究する。テストを受ける。最高の可能な得点を得るように努力する。失敗を避ける。
主な動機	学習における成功は、達成可能であると信じる。	罰への恐れと報酬の約束
強調点	学習者と一緒にルーブリックを使う。学習者の自己評価。学習者への文章によるフィードバック。	達成テスト。最終試験。クラス分けテスト。短いサイクルの評価。

「形成的アセスメントとは、教師が自分の進行中の指導手順を調整するために、あるいは学習者たちが自分たちの現在の学習方策 tactics を調整するために、学習者たちの現状についての評価を誘い出す証拠を提供する計画的な過程である。」

また、本稿の冒頭に紹介した OECD 教育研究革新センターの「形成的アセスメントと学力」では、教師と学習者のいずれにもバランスを保とうとして、「形成的アセスメントとは、学習者のニーズを確認し、それに合わせて適切な授業を進めるための、学習者の理解と学力進歩に関する頻繁かつ対話型（インタラクティブ）なアセスメント」<sup>36)</sup>と定義した。

このように形成的アセスメントについては、様々な定義がなされているが、2010年に『形成的アセスメント・ハンドブック』を編集した Andrade, H. L. は、数多くの研究者や実践家の形成的アセスメントの定義を集約して、次のように纏めた<sup>37)</sup>。

「形成的アセスメントは、少なくとも(a)教師や管理職に教師や管理職に学習者たちの学習についての情報を提供して、指導デザインを導き、(b)学びの進歩について学習者にフィードバックし、学習者たちのパフォーマンスと学習目標との間のギャップの縮め方を学習者が決める、という2点において定義できる。」と。

要するに、形成的アセスメント（学習のためのアセスメント）は、教師側と学習者側の両面から、フィードバックの原理である教師がねらった目標と学習者の現在の学びのレベルの間のズレを縮めるための指導法と学習法である。

ところで、アメリカでは、形成的アセスメントをスタンダード準拠のいわゆる大規模テストの成績アップのために役立つのではという期待を持って迎えられた。アメリカ各州で採用されているスタンダードは、問題解決力など高次の思考も盛り込んでいるが、スタンダードに準拠したテストは、現実には、ほとんどが多項選択式による問題で、知識や技能を問うているものが多い<sup>38)</sup>。そして、教師は、「時間がない」という不満を感じており、よりテスト結果に結びつく方策を望む声が根強い<sup>39)</sup>。そこで浮かび上がってきたのが中間的アセスメント interim assessment という方法であるが、「中間的テストとは、ベンチマーチテストのことであり、重要なカリキュラムの結果の習得に関して学習者たちの現状を測定するために定期的に、恐らく2～3カ月に一度実施されるテスト」なのである<sup>40)</sup>。

なお、全米州教育長協議会は、形成的アセスメントと中間的アセスメントを混同しがちであるという問題を解消するために、2008年に次のような中間的アセスメントの定義を行った<sup>41)</sup>。

「中間的アセスメントとは、子どもたち、教室、学校又は学区で政策立案者や教育者に情報提供するために特定のアカデミックな目標に関する学習者の知識や技能を評定 evaluate する通常は授業外で年度内に複数回実施するアセスメントである。」と。

そして、CCSSO は、形成的、中間的、総括的という3種類のアセスメントを表3のように図式化して違いを鮮明にしようとした<sup>42)</sup>。

端的に言えば、形成的アセスメントと中間的アセスメントとの違いは、その範囲 scope にあるが<sup>43)</sup>、Popham は、中間的アセスメントが大規模テストの得点力アップに効果があるか否かに疑問を呈している<sup>44)</sup>。Bloom の診断的評価、形成的評価、総括的評価の3分類がどちらかと言えば教授学習過程に関心を向けていたが、表3のアセスメントの3分類は、それぞれの国や州だけでなく経済開発協力機構の生徒学習到達度調査（PISA）や国際教育到達度評価学会の学力テスト（TIMSS）など国際的な大規模テストをより一層意識していることは間違いない。とは言うものの、アセスメントの語源である“寄り添う”相手が、それぞれのレベルで利害関係者 stakeholders の範囲も広くなったり、狭くなったりするものの、いずれのアセスメントにおいても、利害関係者に対して情報の公開と共有を求めているということである。その点が形成的評価とは大きく違うのである。

#### 4. 形成的アセスメントの到達点

##### 1) 総括的アセスメントとの関連性

Brookhart, S. によれば、形成的アセスメントとしての教師の「コメントが、学習にとって評定以上によいという一般的な結論で決着」した結果、今や「形成的アセスメントと総括的アセスメントをどのように絡めるのかということが主要課題になっている」と言う<sup>45)</sup>。そして、この関連性の問題について、最近、Marzano, R. J. が説得力ある提案をしている。すなわち、形成的アセスメントか総括的アセスメントかということは、その情報の使われ方によって決まるのであり、一連の学習の最後に行うアセスメントには、誤謬を含んでいる可能性もあるので、「総括的な得点は、最後に行う一度限りのアセスメント

表3 形成的アセスメントと中間的アセスメントと総括的アセスメントの比較

	形成的アセスメント	中間的アセスメント	総括的アセスメント
典型的な使用	進行中の教授と学習を調整するためのフィードバック	学習者の進歩のモニタリング	学習者の位置：学校と学区の説明責任
実施の頻度	不断：一日に複数回	一般には年度内に2～4回	通常、年間1回
実施の範囲	学習者と教室	通常、学区	州

ントから引き出されるべきではない」のであり、「総括的な得点は、ある特定の時点での一人の学習者の最終的な現状について、最も理に叶っており、その現状を代表するものでなければならない」のであって<sup>46)</sup>、そのために、次のような原則にしたがうべきであると言う<sup>47)</sup>。

- (1) 形成的アセスメントは、(ア)テストや演示のように授業を中断して行う、(イ)観察のように授業中に行う、(ウ)学習者が「これを見て欲しい」と申し出て行うという3種類がある。
- (2) 形成的なアセスメントは、得点化され、それを記録しなければならない。ただし、指導としてフィードバックをする際には、得点化することもあれば、しないこともある。
- (3) 特定の学習者の長期に渡る進歩を辿るために、記録した一連の形成的な得点を使う。なお、指導のためのフィードバックでは、全体や個々の学習者の現状を把握して行う。
- (4) 一連の形成的な得点は、総括的な得点を出すために使う。指導のためのフィードバック情報は、総括的な得点をデザインする際には使わないが、総括的な得点を決めるために、役立つようなら、使用することもある。

実は、私もかつて中学校教員との協働研究でこのような評定を行ったことがあるが<sup>48)</sup>、Marzanoが優れているのは、内容のレベルを①事実情報であるが、トピックに重要な多項選択式問題、②学習者が原理を説明したり、一般化の例を示したりする問題、③学習者に推論させたり、授業でやった事柄を越える応用問題に分けて、より詳細な評定の方法を示したことである。そして、これらの形成的な得点を総括的なアセスメントに生かすために、④最後の評定時に得点の変容をみる、⑤得点をその都度積み上げていく、⑥学級全体の進歩をみる、⑦1年間を通じた継続的な改善をみるという方式の使い分けを提案した<sup>49)</sup>。

## 2) フィードバックの手立て

Hattie, J. は、800以上のメタ分析をして、子どもの達成への影響要因として最大のものがフィードバックであるということを明らかにした。彼の著書『目に見える学習』によれば、教師か

ら子どもに与えるものがフィードバックであると思っていたが、子どもから教師に与えるフィードバックのほうが強力であることが分かったと述懐しているけれども<sup>50)</sup>、その発想は、すでにSadlerの前掲論文の中に込められていた。このように子どもから教師への評価情報の提供が従来の形成的評価との違いのひとつである。

さて、フィードバックの活かし方については、Brookhartがその方略についてタイミング、量、様式、発表相手の4つに分けて、良い例と悪い例を示しながら、詳述した。また、フィードバックの内容については、焦点、比較、機能、バランス、明瞭性、明細性、トーンに分けて具体的な方策を示している<sup>51)</sup>。しかも、学習者自身がフィードバックを活用する方法も提案し<sup>52)</sup>、成績の良い者や学習困難な者など多様な学習者のニーズに応じたフィードバックの方法も提案している<sup>53)</sup>。

## 3) 指導の枠組みにおける位置づけ

形成的アセスメントの導入モデルとして、Blackたちは、教師が学習者の声を聞いて、解釈するという形成的アセスメントの要点については社会構成主義に基づく自己調整学習を、目標とフィードバックについては認知的加速プログラムやダイナミック・アセスメントを推奨しているが<sup>54)</sup>、これらは、すでにわが国においても検証の試みがある<sup>55)</sup>。

他方、アメリカに目を転じると、形成的アセスメントを指導の枠組みに位置付ける方策として2方向からの提案がある。第一には、Pophamの学習進行 learning progression モデルである。これは、高校歴史の授業では、①歴史の教訓、②歴史的事件の理解(多項選択式問題)、③関連する歴史の教訓の確認とか分析的な小論文づくりや、高次の認知技能のような①カリキュラムのねらいとそれを達成するための情報や原理や手順と、それを生み出す②生きて働く知識 a body of enabling knowledge、③特定の下位技能の二つで構成される「建築ブロック」と称する基本要素を配列したものであると言う<sup>56)</sup>。この考え方は、Williamもその必要性を認めているが<sup>57)</sup>、どちらかと言えば中等学校向けであろう。



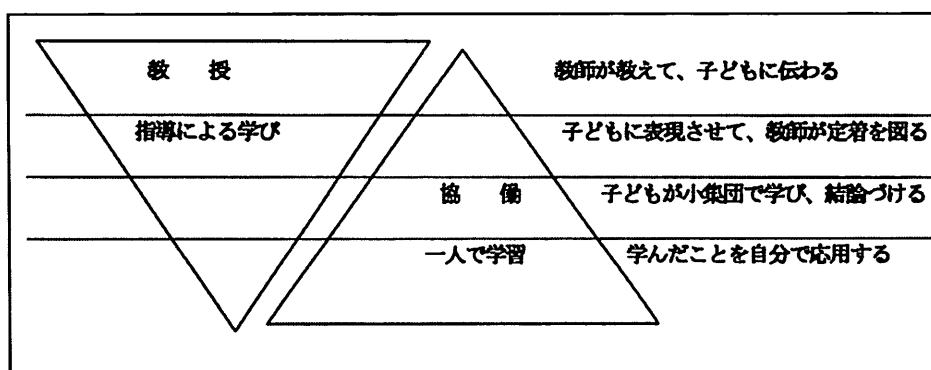


図1 責任移譲モデル（教師主導の指導から子ども主導の学習へ）

もう一つの注目すべき指導的枠組みは、Fisher, D. と Frey, N. が提案するものであり、その基盤にあるのは、上図のような4つの学習形態に着目して、Sadler がかつて述べたように、教師主導から学習者主導に学びの所有権と責任を移行させる考え方であり<sup>58)</sup>、年齢相応の責任を果たさせるための手立ても示している<sup>59)</sup>。また、目標設定に関しては、知識の習得とパフォーマンスに分けて、テストやチェックリストだけでなく要約的な書き方やプレゼンテーションなどの活動別のルーブリックも示している<sup>60)</sup>。Frey は、特別支援教育の専門家でもあるので、言語活動を軸にした授業への心配りも手厚い。このような意味で、彼らの提案は、我が国の教育状況にも合致した貴重なものではないかと思う。

## 5. 結びにかえて

形成的アセスメントは、「教師は、教える人、子どもは学ぶ人」という固定観念を打破する方法論になりうる。それは、目標から内容を引き出し、方法を適用して、最後に評価するという伝統的な方法に対する挑戦でもある。私自身、ポートフォリオ評価や Stiggins たちの評価論に習って約10年前に小学校や中学校の教師と一緒に形成的アセスメント的な「評価と学びの連動」の実践を行い、そこで評価規準の発達に即した適用や良い学習物と悪い学習物を示して検討させる学習機会の提供や活動別ルーブリックを提案し<sup>61)</sup>、また中学生によるポートフォリオ検討会の進め方の実践例も示したが<sup>62)</sup>、理論的枠組

みの中での位置づけ方が弱かった。本稿に紹介したように、形成的アセスメントを指導枠組みの中に位置付けて、具体的にワークシートや教師の指導助言などの提案も示されているので、そのような先行理論や実践から学べば、スムーズな形成的アセスメントの導入ができるのではないだろうか。

英米では、形成的アセスメントの信頼性や妥当性、さらに効果量がどこまで信用できるのかという議論もあるが、最大の課題は、評価科目を必修としていない教員養成教育であり、そのような教育を受けてきた初等中等学校の教師に対する教育であると言う<sup>63)</sup>。我が国でも同様の課題があるが、まずその前に形成的アセスメントに関する地道な教育実践を積み重ね、その成果を共有し、根拠付けることが、これらの課題解決への取組に繋がるのではないかと思う。

## 【引用文献】

- 1) OECD 教育研究革新センター編著（有本昌弘 他訳）「形成的アセスメントと学力」明石書店、2005年、273頁。
- 2) 梶田徹一、「形成的評価のために」、明治図書、1986年、153頁。
- 3) OECD 教育研究革新センター編著（有本昌弘 他訳）、273頁。
- 4) [http://www.ets.org/item\\_bank](http://www.ets.org/item_bank)（2011年6月18日確認）形成的アセスメントは、最も成長するテスト市場と捉えられている。See Cizek, G. J. An Introduction to

- Formative Assessment, *Handbook of Formative Assessment*, edited by Andrade, H. L. and Cizek, G. J., Routledge, 2010, p.12.
- 5) [http://www.michigan.gov/documents/mde/OEAA\\_Fall\\_Conf\\_19\\_Using\\_Interim\\_Assessments2\\_295600\\_7.pdf](http://www.michigan.gov/documents/mde/OEAA_Fall_Conf_19_Using_Interim_Assessments2_295600_7.pdf) (2011年6月18日確認)
- 6) Scriven, M. The Methodology of Evaluation, *Perspectives of Curriculum Evaluation*, edited by Tyler, R. W., Gane, R. M., and Scriven, M., Rand McNally & Company, 1967, p.41, p.44.
- 7) Ibid. p.72.
- 8) B. S. ブルーム他著 (梶田徹一 他訳)『教育評価法ハンドブック』第一法規、1973年、162頁。
- 9) Bloom, B. J. "The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring", *Educational Researcher* 13 (6), 1984.
- 10) Blookhart, S. M. "Expanding Views About Formative Classroom Assessment: A Review of the Literature" in *Formative Classroom Assessment*, edited by McMillan, J. H., Teacher College Press, 2007, p.44.
- 11) Ramaprasad, A. "On the Definition of Feedback" *Behavioral Science*, 28 (1), 1983, p.5.
- 12) Sadler, D. R. Formative Assessment and the Design of Instructional Systems, *Instructional Science*, 18, 1989, pp.123-141.
- 13) Ibid. p.121.
- 14) Ibid. p.143.
- 15) Black, P. and Wiliam, D. Assessment and Classroom Learning, *Assessment in Education*, 5 (1), 1998a, p.8, p.36. 効果量とは、2つのグループの得点差を統制群の標準偏差で割った係数で、その効果量を百分率にも変換して、効果を検証することができる。
- 16) Ibid. (1998a), pp.38-39, p.41, p.45.
- 17) Black, P. and Wiliam, D. Inside the Black Box, *Phi Delta Kappan*, 80 (2), 1998b, p.140.
- 18) Ibid. p.140.
- 19) Brookhart, S. M., op.cit. p.44.
- 20) Black, P. and Wiliam, D. 1998b, pp.143-144.
- 21) Black, P. and Wiliam, D. "Working Inside the Black Box" *Phi Delta Kappan*, 86 (1), 2004, pp.10-16.
- 22) Ibid. p.16, p.19.
- 23) Black, P., Harrison, C., Lee, C., Marshall, B., and Wiliam, D. *Assessment for Learning*, Open University Press, 2003, p.2.
- 24) Ibid. p.122.
- 25) 筆者は、Stiggins たちの理論に学びながら、小学校及び中学校における“評価と学びの連動”の実践をすでに行っている。参照 安藤輝次編著『評価規準と評価基準表を活用した授業実践の方法』黎明書房、2002年。安藤輝次『絶対評価と連動する発展的な学習』黎明書房、2004年。
- 26) Ibid. p.124.
- 27) Assessment Reform Group, *Assessment for Learning-Beyond Inside the Black Box-*, 1999, pp.4-5.
- 28) Ibid. p.7.
- 29) Assessment Reform Group, *Assessment for Learning 10 Principles*, 2002, p.2.
- 30) Black, P. and Wiliam, D., "Developing the Theory of Formative Assessment", *Education, Assessment, Evaluation, and Accountability*, 21 (1), 2009, p.8.
- 31) Wiliam, D., "An Integrative Summary of the Research Literature and Implications for a New Theory of Formative Assessment", *Handbook of Formative Assessment*, edited by Andrade, H.L. and Cizek, G.J. Routledge, 2010, p.24.
- 32) Stiggins, R., Arter J., Chappuis J., Chappuis S., *Classroom Assessment for Student Learning*, Educational Testing Service, 2004, p.33. なお、Black, P. たちの理論に関する研究実践の取り組み的役割を担ってきた Clarke, S. によれば、イングランドにおいては、“形成的アセスメント”は、“学習のためのアセスメント”と同義で使われているが、スコットランドでは、“学習のためのアセスメント”と“総括的アセスメント”のすべてを一緒にした概念として使われていると言う。See Clark, S. *Active Learning Through formative Assessment*, Hoddler Education, 2008, p.9.
- 33) Stiggins, R., "From Formative Assessment to Assessment FOR Learning", *Phi Delta Kappan*, December 2005, p.328.
- 34) Brookhart, S. M. *Exploring Formative Assessment*, Association for Supervision and Curriculum Development, 2009, p.14.
- 35) Popham, W. J. *Transformative Assessment in Action*, Association for Supervision and Curriculum Development, 2011, p.6.
- 36) OECD 教育研究革新センター編著、前掲書、26頁。
- 37) Andrade, H. L. "Summing Up and Moving Forward" *Handbook of Formative Assessment*, edited by Andrade, H. L. and Cizek, G. J., Routledge, 2010, pp.344-345.
- 38) Abraham, M. "Implication of High-Stakes Testing for

- the Use of Formative Classroom Assessment”, *Formative Classroom Assessment*, edited by McMillan, J. H., Teacher College Press, 2007, p.80.
- 39) Ainworth, L. & Viegut, D., *Common Formative Assessment*, Corwin Press, 2006, p.ix.
- 40) Popham, W. J., *Transformative Assessment*, Association for Supervision and Curriculum Development, 2008, p.9.
- 41) TILSA SCASS Interim Assessment Subcommittee, *Interim Assessment Practices and Avenues for State Involvement*, 2008, p.2.
- 42) Ibid. p.4.
- 43) Bambrick-Santoyo, P., *Driven by Data*, Wily & Sons, Inc. 2009, p.127.
- 44) Popham, W. J., op.cit.(2008), p.10.
- 45) Brookhart, S., “Combining Sources of Classroom Achievement Information for Formative and Summative Purpose” *Handbook of Formative Assessment*, edited by Andrade, H. L. and Cizek, G. J., Routledge, 2010, pp.279-280. p.295.
- 46) Marzano, R. J., *Formative Assessment & Standard-Based Grading*, Association for Supervision and Curriculum Development, 2010, p.27.
- 47) Ibid. p.31.
- 48) 安藤輝次、前掲書（2004年）、130-131頁。
- 49) Marzano, R. J., op.cit.(2010), p.40, pp.81-96.
- 50) Hattie, J., *Visible Learning*, Routledge, 2009, p.173.
- 51) Brookhart, S. M., *How to Give Effective Feedback to Your Students*, Association for Supervision and Curriculum Development, 2008, pp.5-7. なお、詳細な手立ては、本書の第3章、第4章、第6章で論述している。
- 52) Ibid., See Chapter 5 (pp.58-75).
- 53) Ibid., p.102.
- 54) Black, P. and Wiliam, D., op.cit.(2009), pp.13-16, pp.18-20.
- 55) 例えば、自己調整学習については、伊藤崇達「自己調整学習の成立過程」北大路書房、2009年、認知的加速プログラムについては、小倉康「科学的リテラシーと科学的探究能力」平成17年度科学研究費補助金特定領域研究報告書、2006年、ダイナミック・アセスメントについては、平田知美「教室におけるダイナミック・アセスメントに関する一考察」『広島大学大学院教育学研究科紀要』第三部、第56号、2007年、などがある。
- 56) Popham, W. J., *Transformative Assessment in Action*, Association for Supervision and Curriculum Development, 2011, pp25-26, p.38.
- 57) Leahy, S. and Wiliam, D., “Devising Learning Progressions” paper presented at the Annual Meeting of the American Educational Research Association held at New Orleans on April 2011.
- 58) Frey, N. and Fisher, D., *The Formative Assessment Action Plan*, Association for Supervision and Curriculum Development, 2011, p.6.
- 59) Ibid., pp.3-6.
- 60) Ibid., p.27, p.41, p.45, pp.74-75.
- 61) 安藤輝次、前掲書（2004年）、144-149頁。安藤輝次編著、前掲書（2002年）、142-144頁。
- 62) 安藤輝次、「学習課題と活動を繋ぐ問いかけを」『教育科学 社会科教育』明治図書、2004年4月号、112-115頁。
- 63) Wiliam, D. op.cit.(2010), p.21. Andrade, H. L., op.cit.(2010), p.348-349.