

今求められる教師の評価リテラシーとは何か

田中耕治

佛教大学教育学部客員教授・京都大学名誉教授

1. はじめに

只今、ご紹介いただきました佛教大学の田中耕治です。本日は、このような機会を与えていただきましてありがとうございます。講演のテーマは「今求められる教師の評価リテラシーとは何か」です。

まず、はじめに「評価リテラシー（英語では、assessment literacy or evaluator competency と申しますが）」とは何かについて、皆さんと共通理解しておきたいと思います。おそらく、教育心理学をご専門とされている研究者の皆さんの多くは、「評価リテラシー」と言いますと、おもに統計的分析手法を想定されると思いますが、私は教育実践を対象とします教育方法学を研究している者ですので、「評価リテラシーとは、教育評価論を機能的、批判的に読み解き、運用（再構成）する能力」とやや広く捉えております（なお、ここで、「機能的、批判的に読み解き」と表現しておりますが、この点に関しましては、「リテラシー」をご専門に研究されている、本日のフォーラムの司会をされている、京都教育大学の樋口とみ子先生にお尋ねいただければと思います）。

このように教育実践のベースで教育評価の研究を本格的に始めましたのは、兵庫教育大学に赴任してからであります。現在の「教職大学院」とは少し違って、当時の兵庫教育大学では、大学院に入学された主に現職教員と一緒に、ある単元に絞って、教育評価を意識した「授業研究」を二年間にわたって行い、それを修士論文にまとめるということをしておりました（その当時の悪戦苦闘の記録は、拙稿(1)をお読みいただきたいと思います）。

本日は、教職大学院での講演ということで、かなり力が入り、盛りだくさんの内容となっております。また講演時間も限られておりますので、あまり横道に逸れずに、エッセンスを押さえつつ、ややはしょった話になりますが、話の内容に興味を持っていただきましたら、挙げております参考文献などでお確かめいただければ有難いです。

2. 教育評価の二つの機能

さて、本論に入っていきます。

まず、そもそも教育評価とは何かについて、お話しします。最近、書きました文章を読みます(2)。

「人づくりという人類固有の普遍的な営みが行われているところでは、その人づくりの成

否を問う方法を編み出し、その程度を判断する必要が生じる。それは、原初的には「イニシエーション」という形態をとり、より洗練された形態として中国の「科挙」制度や中世イエズス会修道士学校の定期試験制度を生み出すなど、各時代各地域において教育評価は様々な形態と機能を生み出してきた」

ここで言う「イニシエーション」とは、その共同体の一員となることを認める通過儀礼のことですね。私の田舎では、「米俵一俵（およそ 60 キログラム）を担ぐことができたなら、一人前だ」と言われておりました。おそらく、このような「イニシエーション」は古くから世界各地で見られるもので、おそらく教育評価の原型になるものと思います。科挙については、その古典として有名な宮崎先生の本（宮崎市定『科挙 中国の試験地獄』中公新書、1963 年）を、またイエズス会つまりジェズイット教団が行った教育の営みについては、E. デュルケームがうまくまとめています（E. デュルケーム、小関藤一郎訳『フランス教育思想史』行路社、1981 年の第二部）ので、ご覧ください。

このように、教育評価は人類に固有の営みだと理解しますと、その営みをやや大胆に概括しますと、二つであったと理解しております。ひとつは、その人づくりの成否を問う社会的文脈で行われた選考や選抜を行う「**証明機能**」とその人づくりの営みの成否を反省的 (reflective) に問う「**指導機能**」です。

近代社会になりますと、ご存知のようにその社会は身分社会であります帰属主義 (ascription) から業績主義 (achievement) に転換して、学歴社会に代表されるメリトクラシー（つまり能力主義）が成立してきます。そうしますと、その業績や能力を測ることが重視され、つまり証明機能が肥大化して指導機能を侵食し始めます。近代社会となった明治日本の初期は「試験の時代」と言われるように、『試験』が多用されますし、その典型は 20 世紀前後に隆盛を誇ります米国の教育測定運動だったと思います。

その教育測定運動を批判して登場してきたのが、「**教育評価 (evaluation)**」概念を創発しました**タイラー (1902-1994)** です。教育評価の証明機能から指導機能へと視座を転換させることで、それこそタイラー原理（教育目標 (objectives) - 授業 - 評価の一貫性と全体性）を提起し、「教育評価とは、教育活動を改善する営み」であることを主張します。このあたり詳しく話したいのですが(3)、時間に制約がありますので、前に進みます。そのタイラーの直系の後継者が**ブルーム (1913-1999)** であり、ブルームの名前は日本ではよく知られていますね。

この「教育評価 evaluation」概念は、戦後日本に紹介されます。ただ、戦後の混乱した状況の中では、タイラーの述べた「教育評価 (evaluation)」概念と彼が批判した「教育測定 (measurement)」概念とが混同されて理解されます(4)（その結果として、次に話します指導要録に教育測定運動が生み出した「**相対評価**」が採用されることになります）。その中にあって、当時の京都大学教育学部に赴任された正木正先生は、そのタイラーの意図を明確に理解されて、次のように述べられています(5)。

「今日周知の如く教育過程において評価（**evaluation**）が重要視されてきている。これは評価という問題が教育過程に新しく付加され、関心されて来たということではない。むしろ、評価は教育過程に融合されている部分であって、従来問題的にも、方法的にも無意識であったものが、新しく自覚されて来たということである。教育活動の自己発展として評価の問題が分節化し、強調されて来たというべきであろう」

私はこの正木先生のご指摘（つまり「評価は本来教育過程に融合されている」ということ）に学んで、先ほど申しましたように「教育評価」という営みを人類史の中に大きく位置づけなおしました。

3. 指導要録（cumulative Record summary—「累加記録摘要」）研究

次に、戦後日本の教育評価のあり方を公的に規定してきた指導要録についてお話しします。指導要録の原語は、cumulative Record summary のことであり、正確に翻訳しますと「**累加記録摘要**」となります。つまり、あくまで摘要（サマリー）なので、その前提に日常で記入する補助簿が想定されております。

しかし、従来、学習指導要領は注目されますが、この指導要録の方はあまり注目されて来ませんでした。文部省の方では、ながく、教育課程審議会の方で学習指導要領を発出して、その後で主に教育心理学の先生たちで構成される「指導要録改善協力者会議」の方で指導要録を検討・発出しておりました。それが、2000年を機に、「教育課程部会 児童生徒の学習評価の在り方に関するワーキンググループ」の方で指導要録を検討することになり、ようやく指導要録も教育課程の一環であることが認められるようになります。

私事で恐縮なんですが、今から三十年ほど前に指導要録の研究を始めました(6)。その時に、各時代の指導要録の作成に係わられた委員のお名前を調査しようとしたのですが、完全に記録に残っていないのですね。文部省にも問い合わせましたが、すぐにはわからないということでした。その委員のお名前全員がわかりましたのは、なんと三十年後で、その成果は、「石井英真・西岡加名恵・田中耕治編著『新指導要録改訂のポイント』日本標準、2019年8月」に所収しています拙稿「指導要録のあゆみとこれから」という論文に掲載しております。この論文の中で、指導要録の分析枠組みを提示しておりますので、興味のある方はご覧いただければと存じます。

表1で、戦後の指導要録の改訂の歴史をまとめております。詳しくは述べられません(7)が、ひとつ興味ある事項を指摘しますと、1971年版指導要録では「所見」欄における「関心」項目が削除されており、戦後の指導要録の歴史の中で、このような情意領域が削除されたことは初めての出来事でした。それはなぜでしょうか。興味のある方はぜひお調べいただきたいと思っております。

表 1 指導要録の改訂史

時期区分	改訂の特徴－教育評価観
第一期 1948年版指導要録	1. 戦前の「考査」への反省と「指導機能」重視－分析評定としての「相対評価」の採用
第二期 1955年版指導要録、1961年版指導要録、1971年版指導要録	2. 総合評定としての「相対評価」の強化と矛盾の激化<1971年版では「所見」欄における「関心」項目の削除>
第三期 1980年版指導要録、1991年版指導要録	3. 「相対評価」の矛盾の「解消」としての「観点別学習状況」<1991年版より登場>の登場
第四期 2001年版指導要録、2010年版指導要録、2019年版指導要録	4. 「相対評価」批判と「目標に準拠した評価」の全面採用、「目標に準拠した評価」と「個人内評価」の結合

このように、戦後の指導要録の改訂は、およそ八回ほど行われましたが、その中で一番大きな転換点は、やはり2001年の改訂であったと思います。その核心は、<およそ50年間続いた「相対評価」を取りやめ、総合評定においても、「目標に準拠した評価」を採用し、「目標に準拠した評価」と「個人内評価」で教育評価観を構成すること>にありました（ここでは、なぜ50年間の長きにわたって「相対評価」が採用されたのか？という疑問が沸きますが、この点もお考えいただければと思います）。この2001年改訂の基本文書は、教育課程審議会『児童生徒の学習と教育課程の実施状況の在り方について（答申）』<平成12年12月4日>です。その改訂の意義については、以下の五点があげられています（文部省「学習評価に関する資料」平成28年3月14日参照、下線は筆者）。

- <1> 新しい学習指導要領に示された基礎的・基本的な内容の確実な習得を図る観点から学習指導要領に示した内容を確実に習得したかどうかの評価を一層徹底するため
- <2> 児童生徒一人一人の進歩の状況や教科の目標の実現状況を的確に把握し、学習指導の改善に生かすため
- <3> 各学校段階において、児童生徒がその学校段階の目標を実現しているかどうかを評価することにより上級の学校段階の教育との円滑な接続に資するため

〈4〉 新しい学習指導要領では、習熟の程度に応じた指導など、個に応じた指導を一層重視しており、学習集団の編成も多様となることが考えられるため

〈5〉 少子化等により、学年、学級の児童生徒数が減少する中で、評価の客観性や信頼性を確保するため

やや大きな言い方ですが、この 2001 年改訂によって、教育実践を改善する「教え」と「学び」を拓く「教育評価」の本格的な探究がスタートしたと考えて良いでしょう。特に〈3〉に注目してください。「目標に準拠した評価」を採用することで、「高大接続の中に大学入試を位置づける」ことが可能となりました(8)。敷衍すると、「指導機能」から「証明機能」を組み替える展望が拓けたということなると思います。

次に現行の 2019 年版の指導要録の特徴と構造について簡単に説明します。その基本文書は、中央教育審議会初等中等教育分科会教育課程部会「児童生徒の学習評価の在り方について（報告）」(2019 年 1 月 21 日)であり、その冒頭部分に学習評価をめぐる現状の問題状況が次の五点にわたって指摘されています。なかなか耳の痛い内容ですが、ぜひご議論していただきたいと思います。

- 一、学期末や学年末などの事後での評価に終始してしまうことが多く、評価の結果が児童生徒の具体的な学習改善につながっていない
- 二、現行の「関心・意欲・態度」の観点について、挙手の回数や毎時間ノートをとっているかなど、性格や行動面の傾向が一時的に表出された場面を捉える評価であるような誤解が払拭しきれていない
- 三、教師によって評価の方針が異なり、学習改善につなげにくい
- 四、教師が評価のための「記録」に労力を割かれて、指導に注力できない
- 五、相当な労力をかけて記述した指導要録が、次の学年や学校段階において十分に活用されていない

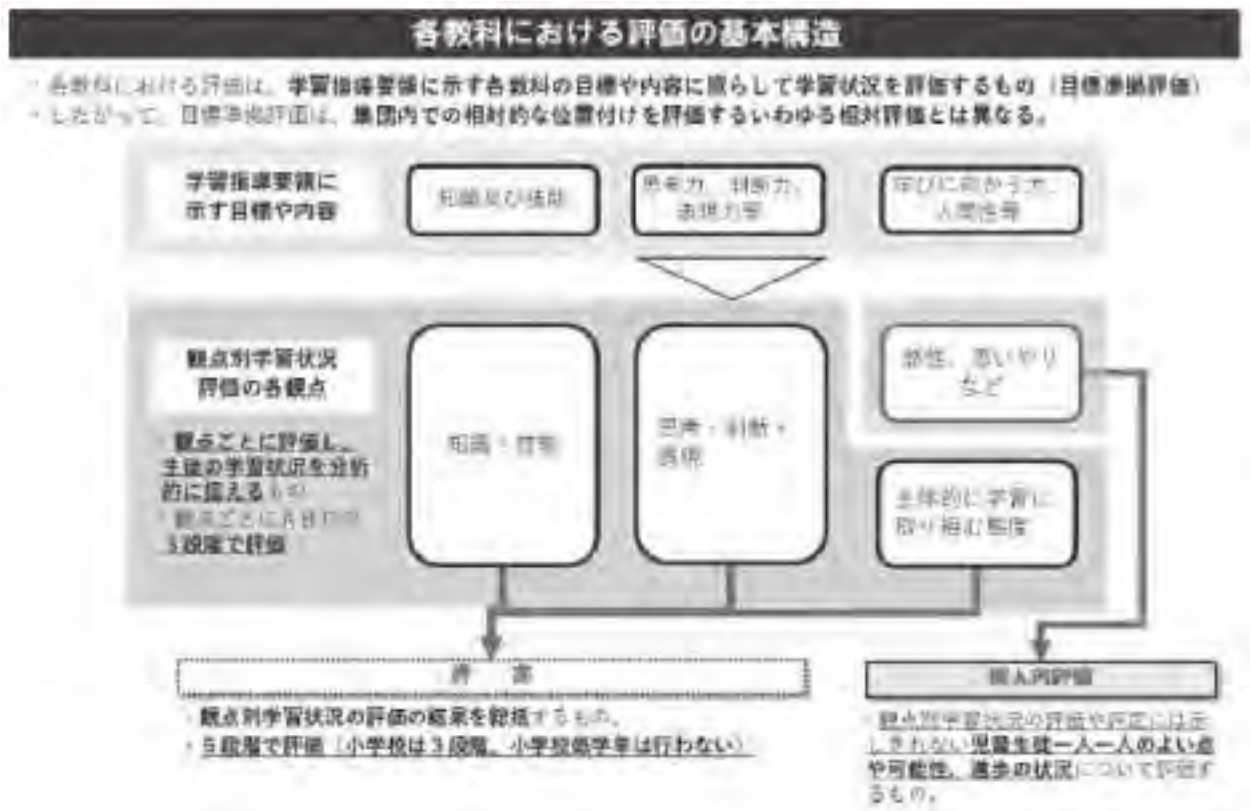
このように指摘した上で、今後の方針として、次の三点が示されています。

- 【1】児童生徒の学習改善につながるものにしていくこと
- 【2】教師の指導改善につながるものにしていくこと
- 【3】これまで慣行として行われてきたことでも、必要性・妥当性が認められないものは見直していくこと

なお、現行の指導要録の構造や機能については、図 1 のように、国立教育政策研究所の方が『学習評価の在り方ハンドブック小・中学校編』2019 年における的確にまとめられていますの

をご覧ください。

図1 学習評価の基本構造



おそらく、教育現場では、観点別評価における、いわゆる「CCA」問題あるいは「AAC問題」、さらには観点別評価と総合評定の関係はどうあるべきかが話題になっていることと思います。この点についても議論いただきたいと考えております。

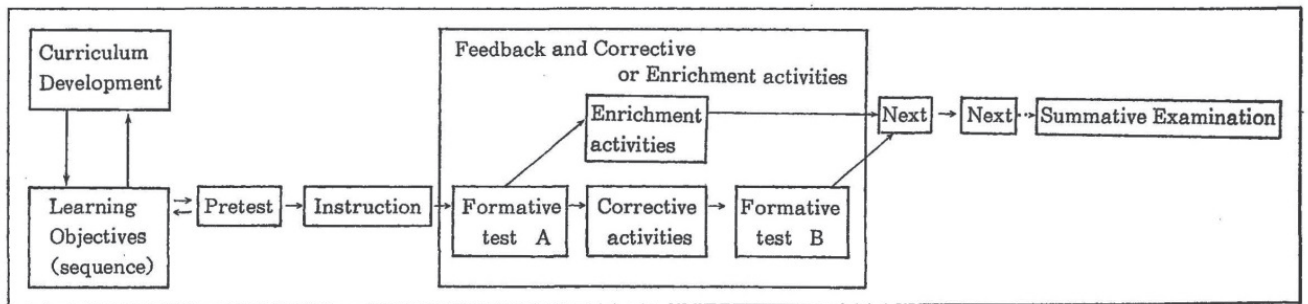
4. 教育評価研究の新しいステージ

指導要録が2001年に改訂される頃に、教育評価研究は新たなステージに立って展開されたように思います。今回は、「形成的評価」の展開と「真正の評価」論の提案の二つに絞ってお話します。

(1) 「形成的評価」から「形成的アセスメント」へ

皆さん、よくご存じのように授業において「形成的評価」を実施するマスタリー・ラーニングを提唱したのは、ブルームでした。そのマスタリー・ラーニングの構造図は、ブルームの高弟であるガスキーがまとめたものを示してあります（図2参照）。とくに、マイノリティの学習権保障にあって、授業のプロセスの中で評価を行うという提起は、当時は大変に新鮮で魅力的に映りましたし、今や「形成的評価」は、日本の授業実践においてもポピュラーなものになりました。

図2 マスタリー・ラーニングの構造図



出典：Guskey, T.R. (1985). *Implementing Mastery Learning*. Belmont, CA: Wadsworth, p.9.
 参考文献：梶田叡一『形成的な評価のために』明治図書、1986年

私は、兵教大に赴任して以来、現職のゼミ生たちと一緒に「形成的評価を活かした授業づくり」に邁進しておりました。1994年にまとめました小文「形成的評価を生かした授業改革論」『現代教育科学』No.447、1994年2月号の中で、「形成的評価」論に関する当時の私の考えをまとめております。以下三点です。

まず、第一点は、「形成的評価」の日本における源流は、日本の「授業研究 (Jugyo kenkyu)」をリードしました東井義雄先生や斎藤喜博先生たちの実践にあるということです。たとえば、斎藤喜博先生の「〇〇ちゃん式まちがい」(9)や東井義雄先生の「つまずきを活かす思想」(10)です。このように、日本の「授業研究」に探りを入れますと、何も「形成的評価」は小テストに限定する必要はなく、日本の教師たちが実践している「発問、机間指導、ノート指導など」も「形成的評価」であり、そのように位置づけても良いと考えるようになりました。

二点目は、「形成的評価」は授業中に無闇にたくさん実施するのではなく、「形成的評価」を行うポイントがある。そのポイントは、その単元の重要事項や子どもたちがつまずきやすいところである。そのポイントを確定するために、ゼミ生たちと先行実践（たとえば、「割合」単元や「ごんぎつね」教材などに関して今までの教育ジャーナルに発表された実践記録）と過去に実施された学力調査を徹底的に蒐集・分類・分析する作業を行いました。

この作業は、教育評価論の用語では、「**内容妥当性**（その課題が単元で求められているねらいを正確に反映しているのかの専門的判断）」を追究する行為であり、「形成的評価」実践の一つ目の重要点ですね。

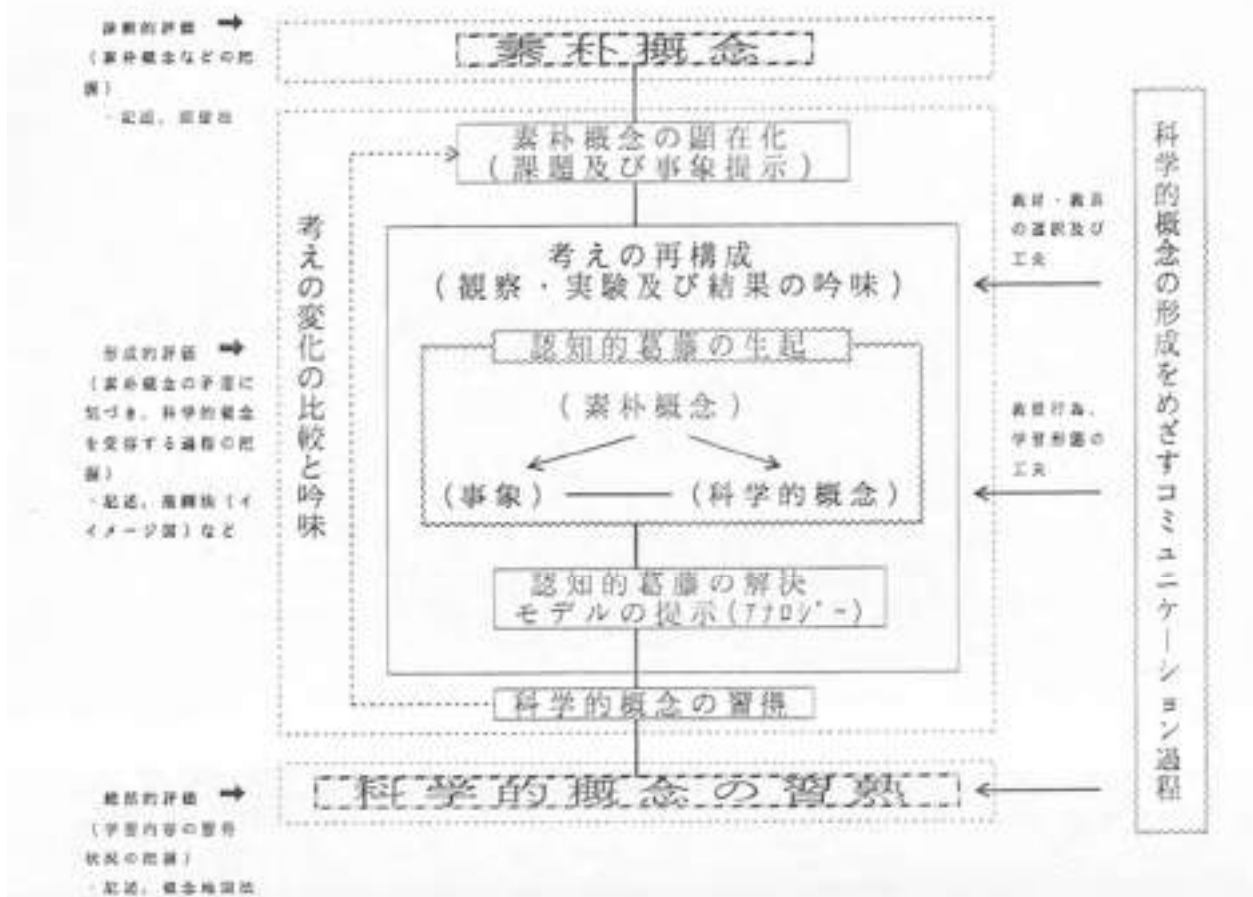
三点目として、ブルームの教え子たち（ガスキーやブロックたち）が「マスタリー・ラーニングとプログラム学習」との異同を探究している理論活動に学んで、むしろブルームの「形成的評価」論にあった行動主義的な学習観（またはプログラム学習）の残滓に自覚的になったということです。

『現代教育科学』に小文を発表しました直後の1994年の春に、ゼミに入学してきた宮崎県出身の現職の院生と出会い、「**素朴概念 (naïve concept)**」研究から構成主義的な学習観を学

び、自己評価の重要性を自覚するようになります(図3にあります、「素朴概念」を念頭に置いた授業の基本構造をご覧ください。とくに「形成的評価」は子どもたちの認識の変容を自覚させる「自己評価」として位置づけられていることに注目ください)。

ちなみに、オズボーン&P. フライバーク編、森本信也&堀哲夫訳『子ども達はいかに科学理論を構成するか』東洋館出版社、1988年が翻訳されて、「素朴概念」研究が盛んになったことはご存じのことと思います。少し話が飛びますが、京都大学を退職して、2018年に佛教大学にお世話になりました時に、理科教育ご担当の平田豊誠先生とお会いして、問題意識が同じであることがわかり、感激いたしました。その後、平田先生から、平田豊誠他「児童・生徒の電流概念の保持状況—小学2年生から高校2年生を対象とした縦断的な調査」『佛教大学教育学部学会紀要』21号、2021年9月という論文をいただき、大変に勉強になりました。すなわち素朴概念の発達論的検討ですね。

図3 授業の基本構想



出典：拙著『学力評価論入門』法政出版、1996年6月の第3章「形成的評価と授業の設計」、p.135

参考文献：拙稿「『学び』をとらえる形成的評価の試み—『素朴概念』に着目した授業づくり(小学校4年生理科『電気や光のはたらき』)に即して—」吉本均代表『教育方法学研究における「知の枠組み」(パラダイム)に関する学際的・総合的研究—戦後授業観の総括と21世紀教育への展望—』科学研究費補助金研究成果最終報告、1996年3月所収。

兵教大で、「形成的評価」の実践的研究を行っていました頃に、英米圏の方で、同じく「形成的評価」の捉え直しが始まっていたということを、京大のゼミ生から学びました。そのことを簡単にお話しします。

まず、ひとつは、「形成的評価の目的ですね。何を目的として形成的評価を行うのか」という点での見直しです。ブルームは、「形成的評価」を説明する時に、「診断的評価」は授業前、「形成的評価」は授業中、「総括的評価」は授業後に行うという風に、その実施する「時期」に応じて提起しました。1980年代の中頃から、アメリカやイギリスでは国家規模の学力向上政策とそれを支えるための強力なテスト政策が実施され、「形成的評価」の名のもとに子どもたちの出来不出来を点検する道具になっているという現状がありました。つまり、最初に申しましたが、評価の証明機能が指導機能を侵食し始めていたのです。そこで、実施する「時期」ではなく、「形成的評価」の「目的」を明確にする必要が生まれ、「学習の評価 (assessment of learning)」「**学習のための評価 (assessment for learning)**」「学習としての評価 (assessment as learning)」という区分が提案されます。表2では、京大の石井先生が、その違いをうまく整理されています。つまり、「学習の評価」とは「証明機能」が目的であり、「学習のための評価」とはまさに「指導機能」が目的ということですね。「学習としての評価」も「指導機能」が目的であることには変わりなく、さらに踏みこんで、子どもたちの「自己調整」を含んでいることが特徴になります。

表2 学習評価の目的

アプローチ	目的	準拠点	主な評価者
学習の評価	席次、進級、卒業証書などに関する判断	他の生徒	教師
学習のための評価	教師の授業に関する決定のための情報	外的なスタンダードや期待	教師
学習としての評価	自己の学習のモニターおよび、自己修正または自己調整	個人的な目的や外的スタンダード	生徒

出典：L. M. Earl, *Assessment as Learning: Using Classroom Assessment to Maximize Student Learning*, Thousand Oaks, CA: Corwin Press, 2003, p.26.

出典：石井英真『現代アメリカにおける学力形成論の展開 増補版』東信堂、2015年、p.324より

まさしく、「学習のための評価」（さらには「学習としての評価」）としての「形成的評価」が明確にされるのです。表3では、同じく石井先生が「形成的評価」概念の拡張を誰によってどのようになされたのかをわかりやすくまとめています。「生徒が自身のパフォーマンスの改善に使えるもの」「生徒を動機づけるもの」という機能が付加されている点に注目してください。そして、ブルームの「形成的評価」論と区別するために、それらは「**形成的アセスメント**」論と総称されるようになります。ちなみに「assess」の語源は「〈子どもの〉そばに座る」という意味です。