

Test] となったことは象徴的な出来事であった。まさしく、大学入試が「教育評価」の一環として把握し直され始めたのである。ニコラス・レブソ著、久野温徳訳「ビツグ・テスト」早川書房、2001年(原著1999年)と荒井克弘・橋本昭彦編「高校と大学の接続——入試選抜から教育接続へ」玉川大学出版部、2005年を参照。

第2章

「教育評価」の混迷と再生

第1節 「エバリュエーション」と「教育評価」のあいだ

第二次世界大戦直後の日本において、アメリカの進歩主義の考え方とともに、「エバリュエーション」も紹介され、「教育評価」と翻訳されることになった。そして、この時期に『教育評価』の題名を持つ著作が多く出版されることになった。しかしながら、この「翻案」事情は、それほど単純なものではなかった(図2-1参照)。

そのひとつには、「エバリュエーション」概念の発信源となったアメリカでは、その創始者であるタイラーの理論と、マジヤメント派のある人々(修正測定派)がその理論を併呑・修正した理論が、同じく「エバリュエーション」を呼称していたことがあった。事実、この時期に日本に紹介された文献には、この二つの系譜があったり、そのために、「エバリュエーション」と「マジヤメント」の区別が判然としない中で、「教育評価」が語られるようになる。

そして、この事態を助長した日本側の条件として、その概念の紹介を担った主体が、おもに「測定(マジヤメント)」概念の研究を戦前に行っていた教育心理学者とその門下生にあたる人たちであったことである。ただし、それらの教育心理学者も一様ではなく、その立場の相違が翻訳に際しても反映されることになった。また、「エバリュエーション」概念が導入されようとした日本の土壌は、明治以降からの「考査」概念が支配的であった。これらの事情も、その紹介に際して独特の色彩を与えることになった。以下、詳しくみていこう。

(1) 「エバリュエーション」の紹介

ここに波多野完治(1905-2001)の興味深い思い出話がある²⁾。おそらくは第二次大戦直後の頃であろうか、波多野によれば、日本では「評価」ということは、授業の結果が子どもにどのように現れたかを「ネグミ」することの意味で

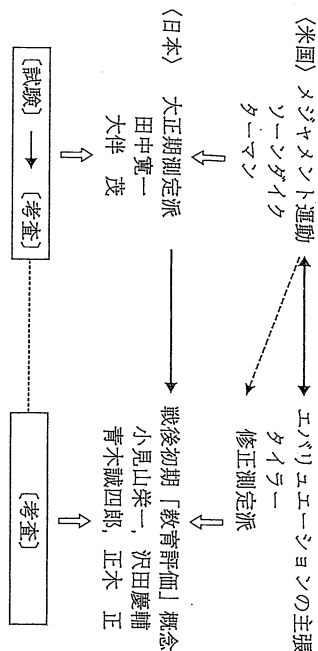


図2-1 日米の理論構図

理解されていた。しかし、そのような意味で外国人と会話するといつも話がずれるという奇妙な経験をする。そして、彼らが「評価」というのは、実は授業で使用する教材の当否やその使用のされ方つまり教師の側の活動を評価対象として語っていることがわかったと述懐している。戦前の日本においては、教科書は国定であり、そのもとでの「指導法」も定型化されているという状況では、その規準に合わせて子どもたちを「評価=ネズミ」(当時の用語では「考査」)するという結果となっていたのである。

すでにみたように「エバリユエーション」概念としての教育評価の考え方が成立するのは、1930年代初期のタイラーの活動に負うところが大きい。ここでは、「エバリユエーション」の目的は、子どもたちを「ネズミ」することではなく、そのような評価行為を通じて得られた情報をもとにして、教育活動を吟味・改善することにあると考えられていた。

このようなタイラーの「エバリユエーション」概念は、たとえば『学習指導要領一般編』の作成に参画した青木誠四郎(1894-1966)によって次のように受容されていた(青木は当時の用語法であった「考査」で説明している)⁹⁾。すなわち、過去の考査法の反省から、「ただ子供を試してみても、子供がどれだけ学んだかの評価をするというようなものでもないし、子供の成績の点つけのためにするのでもないし、まして、子供に脅しをかけて勉強させるためのものである」とした上で、「考査はむしろ先生が自分の指導を反省して、これからの指導の完全を期するためにするものである」と規定する。評価行為の目的はまず何よりも指導の反省の契機として理解されたのである。また、長島貞

表2-1 「教育測定」と「教育評価」

「教育測定」	「教育評価」
学習効果を客観的非人間的操作によって厳密、かつ正確に量的に表現しようとする。用具の客観性、信頼度と統計的基準を重視。主として、教科の達成と知識と、児童の孤立的な行動に關係する。	価値に関連し、流動する教育目標に向かう児童の成長及び発達過程に關心を持つ。用具の妥当性を本質的なものと考え、児童の経験と学習の動因と全体的行動に關係する。

出典：長島貞夫「エバリユエーション」『新教育事典』平凡社、1949年、22-28頁より作成。

夫(1917-1988)もタイラーを引証しつつ、「教育評価」概念と「教育測定」概念の相違を明確にして、「教育評価」の意義を強調している(表2-1参照)。

そして、文部省が作成した教職教養のための教科書『初等教育の原理』(1951年6月)で、このような「教育評価」の意味内容はさらに明確に規定されることになる⁴⁾。そこには、戦後初期「教育評価」概念の特徴が次の5点にわたって説明されている。当時の「教育評価」論の高い水準を知る上で貴重な証言である。①評価は、児童の生活全体を問題にし、その発展をはかるうとするものである。②評価は、教育の結果ばかりでなく、その過程を重視するものである。③評価は、教師の行う評価ばかりでなく児童の自己評価をも大事なものである。④評価は、その結果をいつそう適切な教材の選択や、学習指導法の改善に利用し役だてるためにも行われる。⑤評価は、学習活動を有効ならしめる上に欠くべからざるものである。

このような説明を読むかぎりでは、今日の教育評価論はその上に何をつけ加えることができるかと自問せざるをえないほどに、その説明水準の高さに瞠目させられる。しかしながら、このような「教育評価」と「教育測定」を区別して論じ、前者の意義を強調する意見は、その後、徐々に「修正」されて、「測定と評価」の調和的理解がはかられるようになる。たとえば、戦前の「教育測定」論のリーダーであった田中寛一(1882-1962)が『測定と評価』誌の創刊号(1951年1月)に寄せた「創刊に際して」によれば、最近の「評価」で問題にされていることはすでに「測定の原理」で考えてきたことであり、本誌にあえて「評価」ということは用いなくてもよい」という意見を開陳している。それは、ある意味で「測定」論からの巻き返しの宣言を象徴するものであった。

(2) 「教育評価」概念の融解

(a) 「教育測定」と「教育評価」の関係

この調和的理解をもっとも明快に示してくれているのが、増田幸一(1988-1982)の所論である⁹⁾。まず、最近の「評価によって測定が駆逐されてしまう」という理解は、まったくの誤解であると指摘した上で、両者の関係を「発展関係」と「包摂関係」の二つの関係で説明する。「発展関係」とは、アメリカにおいては「試験→測定(テスト)→評価」という図式で発展してきたという事実を承認することをいう。しかし、ここで増田は、次の点に留意せよという。つまり、このような「発展関係」で「測定と評価」の関係をみると、「それはわが国では評価を強調する結果、測定を過当に軽視することになる。日本では、「測定」が未発達であるから、「測定」と「評価」を同時に発達させなくてはならない。つまり、増田の主張では、日本においては事実上「測定と評価」の「発展関係」は妥当せず、むしろ共存すると理解すべきなのである。

次に、「包摂関係」とは、この両者の共存の関係を具体的に指摘したもので、「評価は測定を内蔵したより広い概念」として理解される。つまり、「教育測定」と「教育評価」の関係を評価対象の拡大として解釈したのである。そして、「測定」以外で「評価」に含まれるものとして、「行動・興味・要求・適応などの実相を捕える活動」と「学級・学校の経営や地域社会の教育運営」があるとされる。このように「測定と評価」を「包摂関係」でとらえることによって、一方では「知能的発達及び学習上の成果」は従来の「測定」論の射程の中で把握すること(「知能検査」や「正規分布曲線」などの使用)を可能にし、他方では「行動・興味・要求・適応などの実相を捕える活動」においてはその評価方法の(教育測定派が要求する)客観性・信頼性を確保するという課題を自らに課すことになる。事実、このような解釈に基づいて、戦後初期に発刊された『教育評価』という題名を持つ著作の多くは、測定・評価の方法に関する記述で大部分が占められるようになる。そして、このような解釈をアメリカの修正測定派に属する人たちの文獻が支えたのである(コラム2-1参照)。

このように戦後初期「教育評価」概念は、それを紹介した人々の来歴によって屈折され、限りなく「教育測定」概念に近い解釈が行われることになる。そ

—コラム2-1 「国語科の評価」と「国語科における評価」—

ここに、戦後初期の国語科学習指導要領作成時の興味深いエピソードを興水実(1908-1986)が伝えている。興水によれば、その作成時において、「国語科の評価」と「国語科における評価」のどちらを選択するかで議論になったと言う。後者は、「生徒の学力の評価、進歩の評価」を意味しており、中学校高等学校の委員会では議論の末に前者を選択することになった。その理由は以下のようなことであった。

「国語科の評価にしたのは、評価は生徒の学習の進歩を評価するだけでなく、そのことを通して、教え方のよしあし、教育計画のよしあしも評価すべきで、「国語科の評価」ということが、国語科の教育課程の評価と取られてもいいし、そういうことがむしろ必要ではないのかという意見が強くなったからであった」。にもかかわらず、小学校ではなぜ「国語科における評価」が当時選択されたのかは疑問の残るところではあるが、以上の議論には明らかに「評価」(国語科の評価)と「評定」(国語科における評価)の概念上の区別が自覚的に語られていたということが大切である。すなわち、「教育評価」とは文字通り「教育活動を評価する行為」であって、そのための重要なデータとしてテストなどの「評定」の結果が活用されるという関係なのである。

しかしながら、このエピソードに再度目を向けると、「評価」がその後にあゆむ運命をある程度予感できるように思われる。すなわち、明らかに概念上は区別されるべき「国語科の評価」と「国語科における評価」に、同じく「評価」という用語が充たされている。この概念と用語の未分化は偶然に生じたものではなく、その背景には「エバリュエーション」が「教育評価」と邦訳される際に動いた「メジヤメント」パラダイムの強い磁場が看取できよう。

出典：興水実「国語科における診断と評価の問題点」(初出1976年)、『国語教育基本論文集』第30巻、明治図書、1993年所収。

ここでは、「エバリュエーション」概念に内在していた指導の反省の契機としての評価行為、教育的価値とそれを具現した教育目標との相即関係にある評価行為という観点が後退し、その「教育評価」の課題はメジヤメントのパラダイムに基づく測定・評価法の開発に限りなく収斂していくのである。

(b) 「考査と教育評価」の関係

他方、戦後初期に「教育評価」概念成立に与った人々は、戦前日本の教育評

価値であった「考査」概念をどのように把握していたのだろうか。このことの検討は、戦後初期「教育評価」概念のもうひとつの特質を明確にすることになる。ここでは、沢田慶輔(1909-1995)の所論を中心に考察を加えてみたい⁹⁾。沢田は、タイラーについてのひとつも早い紹介者であるが、一方ではわが国の教育評価史にも論及し、戦後初期「教育評価」概念の集大成の位置を持つ『教育評価事典』(岡部弥太郎監修、国土社、1965年)にその見解をまとめるに至っている。沢田の基本的な主張は、戦後に「教育評価」として主張されている考査方は、すでに戦前の「国民学校」期に提案されていたというものである。

「国民学校」期においては、科目の全体的関連性を考慮し、一時的な試験の方法によることなく、平素の状況を通じて習得、考察、処理、応用、技能、鑑別、実践及び学習態度等を総合的に考査しようとしていた。たしかに、「国民学校」期においては「皇国民錬成の目的に統合帰一」とするという点は批判されなくてはならないが、総合的に考査することについては「考査法の原理」として好ましい。まさしく、「国民学校」期の考査法が、行動や人格を要素に分解せず全体として評価しようとする主張は、従来の考査法にくらべて、いちじるしく進歩した考査方であって、「教育評価」概念のわが国における淵源であると指摘している。

以上の沢田の歴史理解の根底には、「教育測定」=分析的方法、「教育評価」=総合的方法というシェーマが存在する。しかしながら、はたして戦前の「考査法」と「エバリエーション」としての「教育評価」は同質のものであるのか、詳しい歴史的な論証は本書の第7章にゆずるとして、いわゆる「絶対評価(認定評価)」としての「考査法」と「エバリエーション」を同一視することには、たとえ「学習態度等を総合評定する」という点で表面上の共通点がみられたとしても、まさに「知識と行為の統一(分化したものの再統一)と知行合一(未分化)との間にある決定的な相違」が無視されている¹⁰⁾。厳しく指摘するならば、沢田の「教育評価」解釈は、「エバリエーション」概念を「考査」概念の水準で理解するという問題を内包していたと言えよう(コラム2-2参照)。

以上、戦後初期「教育評価」概念は、紹介者の来歴や翻訳対象となった文献の性格にも影響を受けて、結果として、一方では「測定」概念と他方では「考査」概念と融合して、その「エバリエーション」概念に陥る固有の意義は

—コラム2-2 戦前の考査法の実態—

「考査法の時代」は、「人物第一・学力第二」という国家の教育施策が強化された時代であり、「測定=テスト」が困難な「操行」や学習態度も考査の総合的な対象とされ、結果として教師の主観的な評価を許す結果になった。たとえば、城戸藩太郎(1893-1985)は、その自伝のなかでこの「考査」にかかわる象徴的なエピソードを伝えている。城戸少年は、書や絵が比較的うまくて、いつも「甲」を(甲・乙・丙・丁)の4段階評定で)もらっていた。しかし、ある時、習字の先生が「丙」をつける。その理由として、「おまえは習字はうまいけれど、操行が悪い、だから丙だ」と説明したと言う。どうも城戸少年が理屈っぽい性格であったことが嫌われたとのこと。このエピソードは、「考査」の規準は教師の胸先三寸にあり、もし成績を上げようとするれば、教師に対する人格的な服従や忠誠心を示さなくてはならない「絶対評価」の特徴をよく伝えている。

出典：城戸藩太郎『教育科学七十年』北海道大学図書刊行会、1978年、10-12頁。

未熟の契機に終わろうとしていた。とくに、「測定と評価」の融合は、1960年代になると「相対評価」の脱置という現象に行き着き、波多野の述懐とは逆行して、「評価」=「ネグミ」と受け取られるようになる。この点について考究する前に、当時の教育心理学者のなかでは異質な立場にあった(それゆえに孤高の立場にあった)正木正(1905-1959)の所論を紹介して、日本における「教育評価」概念が持ったもうひとつの可能性に言及しておきたいと思う。

(c) 「教育評価」概念の可能性

先述したように、「教育評価」の任務は、結局のところ測定・評価の方法を開発していくことに収斂していくことになった。しかしながら、言うまでもなく、測定・評価法の開発自体に問題があるのではない。問題の核心は、タイラーも指摘していたように、「どのように測定・評価するのか」の前提に、「何のために何を測定・評価するのか」という教育目標論(または学力モデル論)つまりところ教育的価値論の追究がなされるべき必要があったことである。

この点について、早くから注意を促していたのが正木正だった¹¹⁾。正木は、測定や評価行為を「教育現実の構造や歴史的展開の道」に即して考察すること

で、この問題を「技術の末梢的議論のうちに終止」する傾向を批判する。そして、「教育的価値・目標」を問わないと、評価の管理的・測定的性格は「人間疎外」に陥ると警告するのである。たとえば、「測定基準」とは帰納的経験的に求められる「統計的平均」を意味するが、「評価基準」は、「郷土社会のもつ歴史的社会的規定に立つ基準」であるとして、その相違を明確にしている。

また、正木においてはとりわけ「評価のガイダンス的性格」が強調され、「評価A—目標到達如何の評価」と「評価B—教師自身がAを通して行う主体的な自己評価」が区別されている。この区別は、ひとつの目標への到達をより広い人間形成の立場から意味づけることを要求するとともに、何よりもそのことを行う主体は教師自身であることを強調するねらいがある。とくに後者に関して言えば、戦後初期の「教育評価」概念を唱道した人々の中では、「評価し提供する者」と「教育を改善する者」を分けて、前者こそ自分たちの任務であるということが暗黙裏に前提とされていた。もし、この分離が固定されるならば、教師は最終的な教育主体になりえないということを正木は指摘したのである。

さらには、狭義の測定・評価法に対して、正木はその問題点を明確に指摘している。すなわち、それらの技術の要件として「信頼性」「客観性」「妥当性」をあげ、なかでも「妥当性」こそ対象の「精神構造、成長、教育的指導目標などについての深い洞察」を必要とすることから、もつとも重視すべき条件であるという。この指摘は、「妥当性」を欠いた測定・評価法のひとり歩きを厳しく戒めることにつながる。たとえば、「ゲス・フウ・テスト(guess-who-test)」や「性格検査」は、子どもたちの教育指導の観点から、絶対に使用すべきでないと言っている。そして、「心理学者のテスト」とは異なり、現場教師が容易に取り組める測定・評価法として、「日常のたえざる観察」「組織化された観察」「調査(生活史と実態調査)」をあげている。

以上、正木の提起は、「教育評価」概念の日本における固有な発展の契機を含むものであった。しかし、これらの提起は、当時においては必ずしも主流とはなりえず、正木の提起の内在的な発展は後年に持ち越されることになる⁹⁾。それでは次に、このような出発を遂げた教育評価論のその後の歩みを、その典型的な舞台となった「指導要録」の改訂とそれらに関する論評を手がかりと

して、明らかにしていきたい。

第2節 「相対評価」批判と「到達度評価」の成立

教育評価観とは、評価の規準を何に求めるのかによって区別される評価論の立場をいう。ここでは、戦後日本の教育評価観を規定した「指導要録」の改訂(巻末資料2参照)とそれをめぐる論評を素材として、教育評価観の特質と変容を明らかにしてみたい。それは、未だに終わつた「エバリュエーション」概念を転生させる歴史的経験を追思考することである。そしてこの過程を通じて「相対評価」から「目標に準拠した評価」の歴史的な形態である「到達度評価」へと教育評価観の転換がなされた。

(1) 「相対評価」と「個人内評価」の接合——「加熱」と「冷却」の構造

1948年に文部省は戦後初めての「指導要録」を提起する。その趣旨説明には、「1. 個々の児童について、全体的に、継続的に、その発達経過を記録し、その指導上必要な原簿となるものである。2. 記録事項は、新しい教育の精神からみて、とくに重要と思われるものを選定してある。3. 出来るだけ客観的に、しかも簡単に、かつ容易に記録されるように作られてある。」(文部省学校教育長通達「小学校学籍簿について」)と書かれている。明らかに戦前の「学籍簿」に対する反省を踏まえて、「指導機能」を重視した評価記録簿の作成を促そうとした(1950年に「学籍簿」は「指導要録」に名称変更)。

この上になつて、この「指導要録」は、「学習の記録」欄では「+2. +1. 0. -1. -2」の5段階「相対評価」を、「学習指導上とくに必要と思われる事項」欄と「全体についての指導の経過」欄では記述式の「個人内評価」という。二つの教育評価観を採用している。前者の「相対評価」は、観点が粗い(「社会科」「算数科」では「理解」「態度」「技能」とはいえ分析評定に基づくものであり、さらには教師の主観に従属させられる状態から集団の位置関係を競争によって決めるという状態への変化は、少なくとも「考査」概念に内在していた「絶対評価」からの「解放」を意味していた。また、「学習指導上とくに必要と思われる事項」欄と「全体についての指導の経過」欄に記述式の「個人内評

「評価」が採用されたことは、主観的な規準を絶対化するのではなく、子どもたちを全体的に、継続的に評価しようとしたものである。

たしかに分析評定であっても「相対評価」を導入した点は、「測定と評価」の融合に陥った戦後初期「教育評価」概念に通じる弱点であった。しかし、あくまでも「指導機能」を重視する「指導要録」の立場から、当時文部事務官であり改訂委員でもあった小見山栄一(1913-1963)は、「正規分布曲線」に基づく「相対評価」の絶対化を次のように警告していた¹⁰⁾。①このやり方は、教育測定の精神に則ったものであり、価値的判断を行う教育評価の精神には必ずしも則ったものではない。②この方法では、学年あるいはクラスの差、あるいは学級差を知ることができない。③この方法は、児童の相対的な位置を知ることができるが、教育的にある児童の理解に基づいた処置をすることはできない。たとえば、自分が努力しても、級友も努力したので、成績が上がらなかった場合、その努力を評価してあげられない。④この方法は、統計上の大数の法則に基づいたものであって、わずかに五、六〇人の学級に適用するには無理がある。

しかしながら、次の「指導要録」の改訂(1965年)から、総合評定としての「評定」欄に「相対評価」が位置づけられると、「相対評価」はいわば「ひとり歩き」を始める。この改訂の背景には、やがてすべての国民諸階層を席巻することになる能力主義に基づく学歴獲得競争が、この時期に本格的に始動したことがあげられる。この状況を受けて、「指導要録」の性格は「指導機能」から「証明機能」に転換し、選抜資料として便利であり、かつ「客観性」を保持するという目的で、一教科一評定の総合評定に基づく「相対評価」が採用されることになった。それは、選抜型競争社会の論理が教育世界を併呑していく過程であり、「相対評価」にはこの両者を矛盾なく調停する役割が期待されたのである(コラム2-3参照)。

ここであらためて「相対評価」を規定すると、それは「メジヤメント」運動に淵源を持つ考え方であって、ある集団内での子どもたちの位置・序列(個人差)を明らかにすることを目的として、その「客観性」を保持するために「正規分布曲線」を使用する。通知表にも採用された「5段階相対評価」は有名であり、進路指導で使われることになった「偏差値」とは「50段階相対評価(25点から75点)」のことである(図2-2参照)。まさしく、学力分布が「正規分布」

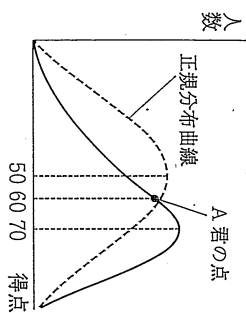
—コラム2-3 「相対評価」の跳梁—

1965年の指導要録は、改訂作業を担った「指導要録研究協議会」における激しい議論の末に提起されたものである。当時神戸市教育研究所長だった黒橋栄一は、この改訂を「指導主義の後退」「学籍簿性格の復活」と厳しく批判している。また、「評定」欄にまさに「相対評価」としての「総合評定」(国語科3や社会科2と評定すること)が採用された事情については、当時の文部事務官の大内茂男が興味深い証言を行っている。「総合評定」の提起の根拠は、①各教科はそれ自体ひとつのまとまりをもって指導されるからであり、②法定公簿はできるだけ簡単な記述が望ましいし、③一教科一評定の方が選抜資料として使いやすいという主張であった。そして、1948年の指導要録にあった「分析評定」を主張する委員(省外委員のほとんど全員)との「妥協」の産物として、「所見」欄が設けられた、というものであった。出典：黒橋栄一「改訂指導要録の批判」、大内茂男「改訂指導要録の「学習の記録」」「児童心理」1966年6月号所収。

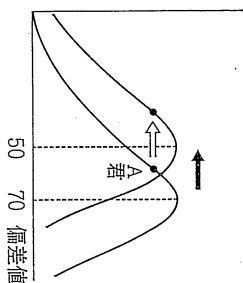
することが「ノーマル＝正常」なのであり、したがって「正規分布曲線」になるようにテスト問題が作成されることになる。

ところで、「相対評価」が跳梁してくると、歴史的な意義を担って登場した「個人内評価」との間に奇妙な関係が生じてくる。すなわち、「相対評価」のもので努力しても成績が向上しない子どもたちに対して、その努力度を「個人内評価」の「所見」欄で「救済」¹¹⁾する(各観点に○印を多く付ける)という関係が機能し始める。

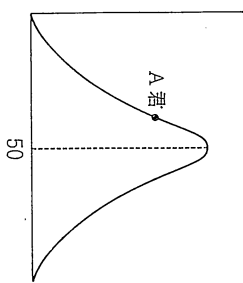
この「評定」欄の「相対評価」の矛盾を「所見」欄の「個人内評価」によって緩衝するという構造、それこそ選抜型競争社会に働いている「加熱(warming-up)」「希少な地位・財産・学歴に向けての競争にカセクト(cathect)させること」と「冷却(cooling-out)」「希望が順挫した場合に、失望・不満・反逆を避けるための機能的必要物——代償的価値に移行したり、改善の達成に満足して適応したりすること」の構造を見事に反映している¹²⁾。その限りににおいて、この教育評価観の二重構造はきわめて強靱な性格を持っている。したがって、「相対評価」を温存して、「個人内評価」を強調しても事態の根本的な解決にはならないばかりか、それでは「個人内評価」それ自体の持つ可能性(後述)をも封殺することにもなる。



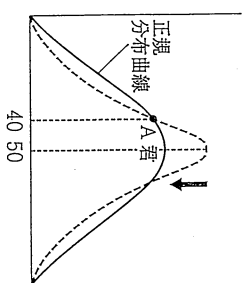
問題が易しかったかあるいは母集団の全員のレベルが高かったために、受験者の得点分布が左のようなカーブを描いたとする。平均点は70点、A君は60点をとったとする。



偏差値は基準が50だからこのカーブを移動させて平均点が50になるようにもっていく。



しかし、平均周辺の人が多かったためにA君のあたりに受験者が密集していて、点差がつかない。



そこで、このカーブを押しつぶして無理やり正規分布曲線にもっていくと、A君の得点60点は偏差値40とでる。つまり、どんな得点分布曲線でもすべて正規分布曲線になおすという考えかたで、実際はこの作業を特定の計算式を使って割り出す。

出典：NHK 取材班『偏差値が日本の未来を支配する』日本放送出版協会、1983年、91頁。
図2-2 偏差値のできるまで

この戦後の選抜型競争社会を下支えしてきた「相対評価」と個人内評価の接合」という評価構造を根本より転換する方向を模索すること、これこそ今日の教育評価観の課題となっている。今後の筆者による論理展開の見通しを明確にするために、やや結論を先取りして述べるならば、その方向性は「目標に

準拠した評価と個人内評価の内在的な結合」であると主張したい。このように筆者が主張するに至った経緯を明らかにするためにも、この問題を歴史の相においてさらに考究してみよう。

(2) 「相対評価」批判の歴史と系譜

「相対評価」に対する批判は、すでに戦後初期から現れていた。たとえば、新庄三郎「ブラダグマイズム教育の批判」(『マルクス・レーニン主義研究』1955年5月)やナガイシヨウゾウ「新しい通信ばつくり」(『教育』1956年3月号)は、その早い時期のものである。そして、1960年代初期には、その批判はかなり明確な形をとってくる。たとえば、中内敏夫「教育評価」(『現代教育学』2、岩波書店、1960年)や滝沢武久「授業と認識」(明治図書、1962年)は、その代表的なものである。

ここでは、『現代教育学』(1961年3月号)に現れた「相対評価」批判の二つの系譜に着目しておきたい。そのひとつは、遠山啓(1909-1979)や桑原作次の立場である。遠山啓は、成績を上げるためには誰かが落ちなくてはならないから、相対評価は排他的な競争を人為的に煽り、「他人の不幸はおのれの幸福」とする心情を形成することになると批判した上で、「もともと教育の目標はすべての子どもが5をとることであり、教師の努力もそこを目標としているはずである。ところが、5段階評価はそれを真つこうから否定していることになる」と述べている。また、桑原作次は、教育という営みは、自然状態にある大多数の傾向である「正規分布曲線」を開いて、「教育曲線」とでもいふべきものに作り替える作業であるから、「教育の結果、依然として正常曲線型を示す」とすれば、それは、問題が適切であったなどと喜ぶよりも、教育そのものが適切であったかどうかについて、むしろ教師の反省を要求しているものとみるべきではないか」と明快に述べている。これらの「相対評価」批判は、明らかにすべての子どもたちの学力保障を願う立場から発せられたものである。

他方、同誌上において東井義雄(1912-1991)や坂本光男は、おもに「生活綴方」の伝統(コラム7-2参照)に接続する立場から、「相対評価」の非教育性を抽象化された点数化にあるとして、子どもたちの内面の成長・発達を記述する方法原理として「個人内評価」を押し出そうとしている。東井義雄は、「一学期

の「3」よりは二学期の「3」には質の高まりがあるはずだ。このちがいを、同じ「3」の中に見よという方が無理ではないか」と述べて、「私は、成績を数字であらわすことには、絶対評価であろうと絶対評価であろうと反対したい」と主張している。ちなみに、大正自由教育の系譜に属する斎藤喜博(1911-1981)もまた、戦前においてすでに「学級全体の横の関係」ではなく、「個人の縦の進歩」こそ重視すべきことを指摘していた¹³⁾。

この「絶対評価」批判の二つの系譜は、敷衍すると同じく「新教育」批判を共通の契機としながら、「教育と科学の結合」原則をもって1960年代の教育内容研究を推進しようとする立場と、「教育と(実)生活の結合」原則を押し立てて教育実践の「生活規定性」に着目しようとする立場の相違を映し出していた。以後具体的に述べることになる「目標に準拠した評価」(その典型的な形態としての「到達度評価」)に対する解釈や異論には、この相違が微妙な影を投げかけることになる¹⁴⁾。

ところで、「絶対評価」の非教育性が論じられ、それに代わって「目標に準拠した評価」(橋本は「絶対評価」と呼称)(コラム1-1参照)への要求が高まっていくなかで、そのような動向を批判して、いわば当時の「指導要録」に理論的な基礎を与えたのが、橋本重治(1908-1992)である。橋本重治の論法は、「絶対評価」「絶対評価」「個人内評価」の長所と短所をバランスよく説明し、とくに「絶対評価」に傾斜しようとする動向を牽制し、「絶対評価」を援護することにある。この橋本重治の論法は、当時の「指導要録」の作成を根拠づけたのみならず、教育評価研究のパラダイムを形成して今日に及ぶ影響を与えることになった¹⁵⁾。

橋本重治によれば、まず「絶対評価」は「教育学」的にみた場合多くの欠点があることを承認する。そこで、「絶対評価」が着目されるが、それは「教育学」的にはすぐれているが、「教育測定学」的にみた場合には主観的で信頼性に乏しい。したがって、外部証明のための機能を持つ「指導要録」には不適格な評価論である。この目的のためには、やはり、「教育測定学」的に「客観性」「信頼性」を持つ「絶対評価」が有効である。しかし、その「教育学」的な欠点があるために、それを補うために「個人内評価」が必要とされる。これが、橋本パラダイムの論法である。そして、この考え方を背景として、1961

年改訂の「指導要録」では、「評定」欄では「絶対評価」,「所見」欄では横断的な「個人内評価」,「進歩の状況」欄では縦断的な「個人内評価」というように配置されていると説明している。ただし、この説明では「絶対評価」の位置が明確でないが、それは、「平常の1回1回のテスト」で採用されているし、さらには、「評定」欄において、「絶対評価に絶対評価を加味する」という形で位置づいているとされる¹⁶⁾。

この橋本パラダイムを、その意図とは逆に「絶対評価」の非教育性を克服して「目標に準拠した評価」に転換するにはどのような課題があるのかという観点から読み込むと、それは橋本が繰り返し指摘する「目標に準拠した評価」としての「絶対評価」の目標設定の主観性に達着する。しかしながら、この点に関して、小笠原英三郎は「教育目標に対する到達度を客観的資料にもとづいて、指導教師が行なう絶対評価をわれわれはもっと信頼してよいと考える。指導要録には、公簿としての客観性に固執するあまり、指導教師の評価はすべて主観的なものという教師不信感が流れている」¹⁷⁾と正鵠を得た指摘を行っている。

この指摘は、言われるところの「絶対評価」の主観性が、けっしてそれに固着した性質ではなく、教師の教育実践の進展に応じて、克服可能なものであることを示した点で重要である。そして、この実践上の課題とは、どの学年で何をどの程度教えるのか(どの程度できたらよいのか)という「到達目標」の設定とそれに基づく測定上の基準の設定ということになるであろう。1960年代を通じての教科教育学は、この「到達目標」の設定という課題に対する理論的・実践的な蓄積を行うことになった。

さて、1969年2月にテレビどのワイドショー番組を舞台に繰り広げられた「通信簿論争」とその延長上に発生した立川二中の「オール3」問題(1972年)は、「絶対評価」問題を考察する上で重要な事件となった¹⁸⁾。その結果、「絶対評価」に対する批判が次の4点に明確に自覚されるようになる。そして、この「絶対評価」への批判を通じて、「到達度評価」が成立することになった。

まず第一の問題点は、必ずできない子がいるということを前提とする非教育的な評価論であること。必ず「1」をつけなくてはならないという「絶対評価」には、いかに指導しようとも必ずできない子どもが存在するという「メジヤメント」運動に発する社会ダーウイズム(優勝劣敗)の思想が流れている。

第二の問題点は排他的な競争を常態化させて、「勉強とは勝ち負け」とする学習観を生み出すこと。成績を上げるためには誰かが落ちなくてはならないから、「相対評価」は排他的な競争を人為的に煽り、「他人の不幸はおれの幸福」とする心情を形成することになる。

そして、第三の問題点は学力の実態を映し出す評価ではないこと。たとえば「5」をもらったとしても、その意味するところは集団における相対的な位置が上位であるということであって、そこで獲得した学力が教師のめざす教育目標に達していたかどうかを証明してはいないということである。むしろ、「勉強とは勝ち負け」であるという学習観が煽動されることと反比例して、「何を勉強したのか」という問いは確実に希薄化していくことになる。これこそ「受験学力」の弊害とされるところである。

第四の問題点として、したがって「相対評価」のもとで学業不振が起ったとして、その責任は子どもたちの努力不足、才能不足に帰せられてしまうことになる。まさしく、「相対評価」は「エバリユエーション」としての教育評価の対極にある考え方なのである。

(3) 「到達度評価」のミニマリスト

以上の1970年を前後する「通信簿論争」を契機として、また、当時アメリカより移入されたグループ学派の影響をも受けつつ(コラム2-4参照)、1970年代の中頃に「目標に準拠した評価」の歴史的な形態である「到達度評価」が名実ともに誕生する¹⁹⁾。ここでは、「到達度評価」の特質を次の4点にまとめて説明したいと思う。

まず「到達度評価」の何よりの特質は、目標論と評価論を表裏の関係としてとらえることによって、「相対評価」を克服するには、その目標論の組み換えを行わなくてはならないと主張したことである。このような発想の原点は、タイラーにその源流を求めることができるが、その日本における初発の契機は勝田守一(1908-1969)の「計測可能」学力説(第4章第2節参照)にあると言ってよいだろう。「到達度評価」のトレーガーのひとり中内敏夫は、この関係認識を「方向目標—相対評価」「到達目標—到達度評価」というシェーマで次のように整理した²⁰⁾。

—コラム2-4 グループの影響—

歴史の不思議な偶然としか考えられないが、1960年代の後半になると、アメリカと日本で「相対評価」批判の狼煙が同時にあがることになった。「エバリユエーション」概念を主張したタイラーの教え子であるグループは、公民権運動に参加することを通じて、社会的に不利な立場にあるマイノリティに「結果の平等」を保障する教育改革の必要性を自覚するようになる。公民権運動の指導者であったキング牧師が凶弾に倒れた同じ年(1968年)、グループの論文「習得のための学習(Learning for Mastery)」が、*The Education Comment*に掲載され、日本に大きな影響を与えた。

この論文のなかでグループは、「学習成績の分布が正規分布に近づけば近づくほど、その教育的努力は不成功であったと断言してよい」と厳しく指摘し、アメリカ社会に根強く存在する社会ダーウィニズムに対して、その「科学」的根拠(同時に「相対評価」の根拠)となっていた「正規分布曲線」神話を論駁した。そして、提唱されたマスタリー・ラーニングは、公民権運動の「夢」を授業改革において実現しようとするものであった。なお、現在のアメリカでは、「正規分布曲線」は、その形状から「ベル・カーブ」とニュートラルに表現されている。

出典：Bloom, B. S., "Learning for Mastery," *The Education Comment*, Vol. 1, No. 2 (1968). B. S. グループ著、稲葉宏雄・大西匡哉監訳「すべての子どもにたしかな学力を」明治図書、1986年(原著1981年)に所収。

〈到達目標〉「2次関数ができる」「江戸時代の産業構造がわかる」というように、何が生徒たちに獲得されなければならないのかを実体的に明示した目標。さらには、その目標と関連する諸目標との間に構造と系統があり、教材・教具を用いて、子どもたちにわかり伝えることが可能なもの。この目標を規準とする評価論が到達度評価。

〈方向目標〉「自然の巧みに関心を持つ」「自ら解決する態度を養う」というように、最低限これだけという限定を持たず、ただ方向を示す方法で設定されている目標。この目標を規準とする評価論が相対評価。

次に大切な点は、このようにして組み換えられた「到達目標」は、「国民の生存権の文化的側面を構成する学習権(教育をうける権利)に対する国家(共同体)の義務」²¹⁾という人権事項のハブスベクティブでとらえられていることである。ここでいう「学習権」とは、「国家主義」や「能力主義」を批判して1970

年代に自覚されてくる教育理念であって、「学習すること」は発達可能態として子どもたちにとってまさに生存の根本にかかわることがらであり、したがって、子どもたちは「わからなければわかるように教えてもらう権利」を有するというものである。この立場から、国の教育行政はもちろん、親の教育権限、教師の研究と教育の自由も、すべての子どもたちの「学習権」保障・学力保障のためにこそ発揮すべきであると主張されるのである。アメリカにおけるブルームが公民権運動に参加して体得したように、人権意識の目覚めこそ「相対評価」を根底的に批判するために必要だったのである。

第三に、「到達目標」の設定が「学習権」保障・学力保障の取り組みを意味するならば、学校における教育課程編成の方法原理は、従来から主張されていた「自主編成」のみならず、「民主編成」が必要との認識が、京都の到達度評価研究の中から生まれてくる²⁹⁾。「民主編成」とは、基本的には学校の教職員集団の民主的な合意に基づいて到達目標づくりを行う営為である。いわば子どもたちの「学習権」保障・学力保障に対する学校の実質的な教育責任(どの教師が担任しようとも、この学校としてはこれだけの学力を子どもたちに形成することを公的に約束すること)を明らかにしようするものである。それに対して、「自主編成」とは、この「到達目標」を達成するために、それぞれの教師が学級や子どもたちの条件に即して、さまざまな授業づくりの工夫を行うことを意味する。

もちろん、この「民主編成」と「自主編成」は機械的に分離されるものではなく、学校の教育課程編成においては、「目標づくり」と「授業づくり」はいわば車の両輪の関係として機能するものである。したがって、たとえ民主的な合意で編成された「到達目標」も、その授業実践の事実(に即して常に改訂の対象となる²⁹⁾)。つまり、このような「民主編成」による「目標づくり」は、目標を絶対化せずに対象化する論理であり、「目標の明確化」とは教育実践に埋め込められている目標を顕在化させ、検討の対象に据えることなのである。換言すれば、「目標の明確化」を否定した上で「学び」を一方向的に強調する論法は、その目標の批判的吟味を等閑視することになるという警告ともなっている。

最後に、もとより「到達度評価」であるかぎり、「目標づくり」に解消されるのではなく、そこで実施される評価方法をも含めて評価対象とするところに

真骨頂がある。一般に、教育実践の本質とは、自らが信ずる教育的価値(具体的には、子ども像や学力像として教師たちによって意識化されているもの)を実現していく、ロマンあふれるホットな行為である。しかし、このことは、時として、その実現過程またはその結果をリアルにクールに診断すること(を)を忘る、「教えたはず」「教えたつもり」という程度の認識で実践が推移することが起こる。すると、このような教育実践は確実に空洞化し始めるのである。「到達度評価」は、目標が明確な評価行為によって、教育実践にリアルズムを貫こうとした。

しかしながら、教育実践をリアルにクールに診断するといっても、それはあくまでも教育的関係性の中で実施されるものであり、その目的は指導と学習に反省的契機を与えることである。そして、この立場をもっとも典型的に示しているのが、「形成的評価」の提唱であらう。従来の教育評価では、実践の最後(のみ)評価を実施していたが(子どもたちの「ネグミ」行為としての評価であれば、これで十分だった)、教育実践を改善しようとするれば、実践の開始時に「診断的评价」を行うとともに、実践過程でも評価行為を機能させる「形成的評価」を実施することが求められる。そして、実践の終わりには学力の総合性・発展性を評価する「総括的评价」が行われるのである。このような評価行為の分化(とりわけ「形成的評価」の提起)とそれに伴う「回復学習・発展学習」の提唱は、「到達度評価」が「エバリュエーション」概念の正統の嫡子であるとともに、その継承・発展を担う立場であることを示した(なお、「形成的評価」の具体的な方法については、第4章第3節で詳述)。

このような特質をもって成立した「到達度評価」は、選抜型競争社会と教育世界を貫いていた「相対評価」システムに有効な批判を加えることによって、自らに学校内改革のみならず評価制度の改革を通じて教育システムひいては社会システムの改革をも射程に入れるという壮大な課題を提示することになった。

ところで、2001年改訂の「指導要録」で採用された「目標に準拠した評価」の実状をみると、あらためて「到達度評価」の歴史的な意義に学ぶことが必要とされている。とりわけ、「相対評価」批判を踏まえない「目標に準拠した評価」が、目標への到達度を単に点検すること、[「相対評価」と同じく子どもたちを「ネグミ」することに陥っているからである。「到達度評価」を踏まえ

た「目標に準拠した評価」であれば、教育実践の改善を通じて子どもたちの学力を保障することがめざされるべきだろう。

その上で、今日の「目標に準拠した評価」は、「到達度評価」の歴史的な制約にも学ぶことによって、概念装置のさらなる精緻化とその創造的な発展を必要としている。それは「到達度評価」の理念を否定することではなく、「目標に準拠した評価と個人内評価の内在的な結合」を実現するためである。次章では、この問題を考察したい。

1) 当時よく読まれた文献は以下のものである。

Tyler, R., "General Statement on Evaluation," *Journal of Educational Research* (1942, 7).

Smith, E. R. and Tyler, R. W., *Appraising and Recording Student Progress*, Harper & Brothers, 1942.

Greene, E. B., Jorgenson, A. N. and Gerberich, J. R., *Measurement and Evaluation in the Elementary School*, Longmans, Green and Co., 1942.

Remmers, H. H. and Gage, N. L., *Educational Measurement and Evaluation*, Harper & Row, 1943.

Ross, C. C., *Measurement in Today's School*, Prentice-Hall, Inc., 1941.

たとえば、グリーンの著作(1942年)は、「序」によれば、「教育テストの使用と解釈(1929年)」、「小学校でのテストの使用と解釈(1935年)の改訂版であり、元来は「教育測定」論として記述されたものである。したがって、そこでの「エバリュエーション(evaluation)」の意味とは、「全体としての子ども」の存在をとりえるためには従来の「紙と鉛筆のテスト」では限界があり、「逐語記録」「インクデビュー」「質問紙法」などの方法を採用することと同義で考えられている。つまり、「エバリュエーション」概念は「測定」概念の単なる方法上の拡大と把握され、そのため本書の大部分は測定・評価法の紹介にあてられている。

2) 渡多野完治「評価の心理学」(1978年7月号)。

3) 青木誠四郎「考査は何のためにするか」(新しい教育) (1948年2月号)、73頁。青木誠四郎については、樋口とみ子「戦後初期学習指導要領における教育評価論——青木誠四郎の場合」田中耕治編「人物で綴る戦後教育評価の歴史」三学出版、2007年所収参照のこと。

4) 文部省「初等教育の原理」東洋館出版、1951年、217-219頁。

5) 増田幸一「エバリュエーション(Evaluation)の意義について」(神戸大学教育学部研究集録) (1955年1月)。

6) 沢田慶輔「教育評価」(講座・学校教育)第6巻、目黒書店、1950年、113-116頁。

7) 中内敬夫「増補学力と評価の理論」国土社、1976年、153頁。

8) 正木正「価値と評価」(教育)第3巻、岩波書店、1952年と正木正「教育の心理」羽田書店、1948年を参照のこと。正木正については、渡辺貴裕「教師の主体性を重視する教育評価論——正木正の場合」田中耕治編「前掲書所収参照のこと」。

9) 正木正の所論を継承した教育心理学者のひとりとして横有恒(1914-1972)がいる。横有恒【教育評価】第一法規、1969年と横有恒【教育心理学の探求】金子書房、1973年を参

照のこと。横有恒については、樋口太郎「理論と実践をつなぐ教育評価論——横有恒の場合」田中耕治編「前掲書所収参照のこと」。

10) 後藤岩男・小見山栄一監修、東京文理科大学内教育心理研究会編「小学校新学修簿の記入法」金子書房、1948年、31-32頁。小見山栄一については、二宮兼一「戦後初期の「教育評価」概念——小見山栄一の場合」田中耕治編「前掲書所収参照のこと」。

11) 「救済」とは、橋本重治のことである(「児童心理」1961年7月号、31頁)。それでは、実際のところはどのように運用されていたのかという点、1955年の指導要録改訂を受け行われた調査(当時愛知県教育委員会指導主事の石川勲の調査【教育心理】1955年2月号、32-35頁)は、興味深い事実を伝えている。すなわち、①「評定」欄で5や4の成績をとっている生徒に○印(所見欄)が集中し、逆に1と2の成績の者には×印が集中するというのである。そして、石川は、これは間違った方法であり、個人内評価の趣旨からは、たとえ5をとった生徒でも×印がつくし、その逆もありうる点と述べているのである。②ある学校を調査した結果、学期間で○印と×印をつける数がいちがうことから、個人内評価では基準や尺度が明確でないために、かなり各教師の思いつきで行われているのではないかと指摘している。この両方の事実は、その後も多くの教育現場で起こっていたことと思われる。

①の結果は、たとえ「相対評価」であっても「1」や「2」を取った生徒に対して、所見欄の観点に○印をつけることは、その観点が学習指導要領と関係していることを考慮すれば、教師にとつてかなり抵抗のあったことを反映している。しかし、石川も指摘するようにこのようなつけ方は「間違った方法」であり、その「救済」を意図するならば、結局のところ②の結果に落ち着くことになる。ともあれ、この「救済」も教師の「温情」から発したものであった。

12) 竹内洋「選抜社会」メディアアクトリー、1988年参照。

13) 松本陽一・高橋嘉明編「斎藤喜博の世界」一笠書房、1983年、310-311頁。

14) 「到達度評価」が、いかなる過去の歴史的遺産に立脚してできたかをめぐっては意見に相違があり、そのことが「到達度評価」を主張する人々にもその力点に相違を生み出すことになる。この点については、さらに厳密な歴史的研究が必要であるが、筆者はひとまず次のように考えている。「到達度評価」は、「教育と科学の結合」原則を主張したといわれる「現代化」運動の成果を立脚点にしているが、その歴史上に成立したのではなく、1960年代後半の「教育の主体性」論争(佐藤興文「現代化における主体的契機の問題」(1968年)、「学力・評価・教育内容」青木書店、1978年所収参照)をくぐり抜ける中で、「教育と(実)生活の結合」原則との統一的理解を展望する時点(【教育】1969年8月号の特集論文参照)に成立したと考える。たとえば、「教育と実生活の結合」は、実は、「科学と教育の結合」を科学主義として排除するところに成立するのではなく、逆にそれは、「教育と科学の結合」の課題と結びつけられて、その問いが、学問(科学)の質の問い直し、「国民化」への志向をうながす契機として理解されるときにはじめて、その正当な意味づけが与えられるのではあるまいか。われわれは、教育と実生活の結合という前に、あるいはそれと同時に、「学問と生活の結合」をいうべきであり、学問と生活の結合とは、学問の国民化と教育の科学化(教育と科学の結合)を同時に含んだものである。この点が、「到達目標」論に影響する(第3章参照)。

15) 橋本重治「相対評価法と絶対評価法」【教育心理】(1961年5月号)、橋本重治については、赤沢早人「戦後教育評価論のパラダイム——橋本重治の場合」田中耕治編「前掲書所収参照のこと」。

16) この「加味論」でいう「絶対評価」は、実際は「個人内評価」の意味である。5段階