

観点別評価の学校教育への導入に関する考察

A Study on the Introduction of Perspective Evaluation into School Education

西 尾 理

NISHIO Osamu

1. はじめに

2002年の学習指導要領改訂以来、小学校、中学校に観点別評価が導入された。その理由のひとつは、「新学力観」による授業方法の転換に沿った学力の定義とその評価方法の転換の産物であった⁽¹⁾。

高等学校でも観点別評価が導入される見通しである。その理由も、変わる兆しが見えない従来の知識偏重の授業方法や「ペーパーテスト偏重」からの脱却を目指すためだという⁽²⁾。

ところが、新学習指導要領の下での学習評価について、中央教育審議会の作業部会（WG）は、観点別評価を現在の4観点から3観点到整理することなどを盛り込んだ報告をまとめた。従来の「関心・意欲・態度」。「思考・判断・表現」、「資料活用技能」、「知識・理解」から「知識・技能」、「思考・判断・表現」、「主体的に学習に取り組む態度」に変更されるという。その理由は、「関心・意欲・態度」が挙手の回数など形式的な行動を評価する誤解があったことや教員の負担軽減緩和のためだという⁽³⁾。

観点別評価においては、導入時には「関心・意欲・態度」、「思考・判断」、「資料活用の技能・表現」、「知識・理解」であったが、「表現」が「資料活用の技能」との組み合わせから抜けて「思考・判断」と組み合わせられるという変更があった。観点別評価の基準となっている言葉（概念）は、抽象度の高い言葉（概念）である。ある意味、普遍的な言葉（概念）であろう。それがここ30年ぐらいの間に組み合わせや観点数が変更になるには、それなりの理由があったのだろうか。教育心理学の学説が変わり、新しい学説が定説となったということなのだろうか。イメージしていた授業方法が変更したということなのだろうか、それとも学校現場との齟齬が生じ、変更せざるをえなかったのだろうか。それにしても迷走というしかないのではないだろうか。ペーパーテストによる知識をどれだけ覚えたかを測る評価方法が、生徒の学力を見る基準として一面的であり、知識偏重の教育方法を促進してしまったという批判は、一定、説得力はあるものの、そのオルタナティブが観点別評価なのであるだろうか。

このような観点別評価は、形成的評価からの影響が見られる。そこで本稿の目的は、こうした形成的評価の影響を受けた観点別評価を学校教育に導入したことへの問題点と課題を明らかにしていくことである。

2. 形成的評価のイメージ

形成的評価とは、梶田叡一によると、もともと形成的評価 (formative evaluation) という言葉を使ったのが、シカゴ大学の哲学者スクリバンで、カリキュラム開発に関する多様な評価を論じる中で、カリキュラムの内的構成をよりよいものとするために、その開発途上において行われる各種の評価活動を総称して形成的評価と呼んだ⁽⁴⁾。その後、ブルームらが形成的評価のための具体的方策として形成的テストやその活用として授業システムを追求したためだという⁽⁵⁾。

形成的評価は、授業のプロセスにおいて実施され、形成的評価の結果はフィードバックされ、授業が狙い通りに展開していないと判断された場合には、授業計画の修正や子どもたちへの回復指導が行われる。成績づけには使われない⁽⁶⁾。したがって、総括的評価のように、行き止まりの評価ではなく、そこから何かが始まっていくような評価。評価することによって実態を捉え、それをバネにして次の学習を、指導を、そしてカリキュラム編成を一層有効なものにしていくのだという⁽⁷⁾。

以上のことから、ペーパーテストで、どれだけ知識を獲得したのかを測るいわゆる総括的評価ではなく、学びに向かう関心、意欲、態度や学びの状態にある思考、資料活用等の技能、判断、表現を評価して生徒のリフレクションのみならず教員のリフレクションによる授業改善も促していく。その結果として知識、理解を含めた学力が向上するというのである。

ところで、かつて読売巨人軍でコーチを務め、9年連続日本一に貢献した牧野茂の著書に『巨人軍かく勝てり V9の秘密』がある⁽⁸⁾。この本の中で巨人軍が強かったのは、単に優秀な選手を集めて個々の能力を発揮させたからではない。チームプレーを重んじたからであることが縷々述べられているが、それを象徴的に表しているのが「巨人軍式人事管理と考課」という項目である。巨人軍の年俵は、チームの勝利にどれだけ貢献したかという「貢献ポイント」で判断される。3割とか20勝とかいう表面上の数字とは違って、プレー上のあらゆるものを総合的にとらえたポイントである。当時、2年連続3冠王になった王が20万ドルプレイヤーになったのは、3冠王になったからではなくて、何百万円のアップに相当する「貢献ポイント」をマークしたからであった⁽⁹⁾。

例えば、投手の評価項目は、①勝敗、②イニング、③コントロール、④スピード、⑤スイフト、⑥変化球、⑦コンビネーション、⑧気力、⑨ヘッドワーク、⑩チームプレー、⑪守備、⑫総合からなり、さらに二つに分かれている。一つは「技術」であり、もう一つは「貢献度」である。これに担当コーチが試合ごとにプラス、マイナスの点数をつけていく。この点数に、監督が修正を加える。これが1年間のトータルの「貢献ポイント」になる。例えば、打者の例として、ある打者は4打数ノーヒットだったが、打球をいずれもバットの芯でとらえていた。この場合「技術」はプラス、しかし「貢献度」はマイナス。よけたバットに当たってヒットになった。これがきっかけでゲームが勝った。これは「技術」がマイナス、「貢献度」ではプラス材料となる。犠牲バンドで打率を下げたが「貢献度」ではプラスになる。さらに、各選手の「貢献ポイント」の基準は、前年度の実績を基準としているので各選手によって「貢献ポイント」が違ってくる。前年度に実績のある投手が完投して3点以内に抑えたら10点満点だが、前年度、実績のない投手が5回を3点以内に

抑えても10点満点になるというように、その時点での個々人の能力からどれだけプラスさせたかで査定するのである⁽¹⁰⁾。つまり、それまで投手なら勝利数や防御率、打者なら打率、本塁打数、打点など最終の結果のみで査定していたが、その過程を査定するようになったといえよう。この査定方式は、その後、各球団に採用されていった。

筆者がこの査定方法を知ったのは今から35年も前のことであったが、教員になり、ブルームの形成的評価というのを勉強した時、真っ先に思い出したのがこの査定方式であった。この査定方式がブルームの形成的評価と類似していると感じている。読売巨人軍は、この査定方式を採用して大きな成果を挙げたが、学校教育において、ブルームの形成的評価を採用すると成果が挙げられるのであろうか。ただし、読売巨人軍では、監督以外にこの査定に各専門のコーチがいて、学校教育の教師のようにほぼ一人で、評価（「査定」）するわけではない。

3. 形成的評価の実際

梶田がこの形成的評価の実践根拠として挙げている学校が、ほぼ附属であるのが気がかりである⁽¹¹⁾。また、次のようにも述べるのである。「『指導のための評価』というテーマでの研究校が、毎時間毎時間、やれ小テストだワークシートだ、観察だ、とデータ集めに精力を使い果たし、指導どころではない、といった状況を呈しているのを実際に数多く見てきただけに、この課題についてもっと真剣に取り組んでほしいというのが率直な願いである」⁽¹²⁾。

したがって、この批判の学校の多くは、公立学校なのであろう。では、「この課題について真剣に取り組めば」解決するのであろうか。形成的評価の特質というべき学習意欲（「関心・態度」）の評価に焦点を絞って検討してみよう。

梶田は、「関心・態度」が大事であるという第一の理由として次のように述べている。「この概念に象徴される主体的で総合的な「構え」の形成が学校教育を通じて追及されなくてはならない。知識や理解、技能を数多くもっていようと、「構え」がないと悪い方向で使われたり、使われないままにさびついてしまったりする。」⁽¹³⁾では、「関心・態度」の目標をどう明確化するのだが、「自然に対する関心・態度」の例示に沿って述べている。「①自然の事物や事象に対して積極的な関心を持ち、②繊細かつ豊かな感性をもって自然に接し、③そこでのさまざまな現われや動きに好奇心や問題意識を持ち、④自分なりに探求し解明しようとして対象にのめりこんでいく」⁽¹⁴⁾。

①について、「積極的に」とはどの程度。何をもってして積極的だというのだろうか。②について、「繊細」とは、「豊かな感性」とは何を基準にそう言えるのだろうか。③について、「好奇心」を持ったとどう判断するのだろうか。「問題意識」と「好奇心」の区別はどう分けて測るのだろうか。④について、対象にのめりこむとはどの時点でどういう基準で判断するのだろうか。教師（たち）や学校全体において、例えば、「積極的」、「繊細」などの基準を予め決めておき、教師の前で個々の生徒たちが現われるパフォーマンスを一瞬にして見てとり、その基準で判断し、評価するということなのだろうか。小学校の児童と違い、中学校、高等学校の生徒は、教師の前で評価が明確になるようなパフォーマンスを常に行ってくれるのだろうか。生徒（たち）の個々の性格の違いをどう考慮すればよい

のだろうか。多くの生徒がいる中で、好奇心を示した生徒を見逃す恐れはないのだろうか。この評価を読売巨人軍と違い、各担当コーチがいて、①～④を専門に見てとってくれたりするわけではない。ほぼ担当の教員一人で判断するのである。

梶田は、基準ではなく、(形成的評価であるから)水準としての枠組みをブルームらの「教育目標のタキソノミー：情意的領域」を頭に入れながら、情意的領域における評価基準設定のためにシンプトム(兆候、現れ、きざし)を挙げてみるということを提唱し、学校で実践研究を行ったという⁽¹⁵⁾。シンプトム(兆候)群が具体的にになれば、それに沿った実態把握の方法を考えるということになるのだという。主要なものとして(1)学年(学期)末の時期における授業中の態度、行動、発言などのチェック項目を準備して、観察する。(2)学年(学期)末の段階で、ノートやワークシートなどの記入内容をチェックする。(3)学年(学期)末の段階で、その期間における学習を全体的に反省させる作文や小レポートを書かせ、その記述内容から問題とする特徴の有無を見てとる。(4)プラスの徴候群、マイナスの徴候群を質問紙の形に構成し、子ども自身に評価させる⁽¹⁶⁾。

上記のような、生徒(たち)の「関心・意識」をノートやワークシートや小レポートのような後付けの“材料”で評価するとしても、他の生徒たちとの公正性を“担保”にした基準をどのように設定していくかには疑問が残るし、教員の膨大な労力がかかることになる。まさに、梶田が数多く見てきた「毎時間毎時間、やれ小テストだワークシートだ、観察だ、とデータ集めに精力を使い果たし、指導どころではない、といった状況を呈している」結果になるのではないだろうか。

ところが梶田は、形成的評価を導入するに際して、次のような前提で論を進めているのである。「柔らかな評価」という言い方で、必ずしも得点化されない評価で、梶田は、次のように言う。「記録として残るわけではない評価、成績づけに繋がっていない評価こそが必要である。」⁽¹⁷⁾さらに、次のようにも言う。評価は何よりもまず客観的で厳密でなくてはならない、という思い込みがまだ根強い。だから「関心・態度」などのような特性は評価すべきではないという意見もあるが、1930年代にタイラーも指摘するように測定でないのだから、評価しても構わない。目標とするところにどの程度近づいたかという評価が大切なのである⁽¹⁸⁾。梶田の主張は、教育評価であって学校教育における評定ではないのであった。したがって、学校内においても対外的にも公正性(という客観性)を求めなくても良いということになる。ただ、ここでの前提となる問題は、この形成的評価を学校現場で導入する際に、この評価と評定の区別がなされていないままとなってしまっていることなのである。

では、次に評価と評定の区別や評価と評定の関係を明らかにしていきたい。

4. 評価と評定の区別

学校現場では、評価と評定があまり区別されず使用されている場合が多い。では、教育評価の研究では、どのように定義されているのか、管見する辞典や事典、文献から整理してみる。

(1) 評価とは何か

まず、評価の定義を行っておこう。広辞苑によると評価とは、善悪・美醜・優劣などの

価値を判じ定めること⁽¹⁹⁾。

手元にある教育学小辞典によると、「教育によって生じた学習、人格、行動などの変化を一定の価値基準によって判定しどのような処置を指導したらよいかを考える一連の過程」とある⁽²⁰⁾。

『現代授業研究大事典』によると、評価とは「教育目標の有効な実現に向けて、すでに行われた教育活動＝指導を教育目標との関連においてチェックし、フィードバック情報を与えることによって、その後の教育活動を調整していくために行われるものである。また指導—評価という目的追求の連環的サイクルの中に位置するものであるため、目標追及する主体がみずからのために行う活動であって他者が行うものではない。」⁽²¹⁾

梶田毅一によると、評価とは、もともと子どもにどの段階から学習を始めさせれば良いのかの認定であり、教育の成果はどの程度のものであるかの確認とを中核とするものであった。これが現在では広義に用いられ、教育活動とそれに関連した各種の実態把握と価値判断の全てが含まれるものとしている⁽²²⁾。

田中耕治によると、評価とは、教育（学習）効果に対する判断行為一般をさす概念⁽²³⁾で、1930年代にアメリカのタイラーによって「測定（measurement）」概念に変わるものとして提唱された。教育活動を評価することによって、教師の指導と子どもたちの学習活動の改善をめざす行為である⁽²⁴⁾。特に、「試験」「考査」「測定」と区別する意図をもって、第2次世界大戦後使用されたものである⁽²⁵⁾。

これら評価の定義から考えられることは、評価がある価値観をもって、子どもの学習の変容のみならず教師の指導改善を目的としたものであるということである。

（2）評定とは何か

管見した限り、教育学では、評定という用語がほとんど使用されていない⁽²⁶⁾。学校現場では、例えば、高等学校の場合、評定は、通知表に示される成績のことで、通知表の学年末に示される学年成績を基に指導要録が作成される。そればかりではなく大学、短大の推薦入試や専門学校の推薦などにも活用されるし、就職の際の資料として使われる。対外的にも重要なものとなる。中学校の場合も内申点などで高等学校の入試で大きな役割を果たしている。小学校の場合も、指導要録を作成する際や中学校への情報を伝える際に大きな役割を果たしているであろう。

こうした対外的な性質を持ち、児童、生徒の進路に関わることから学校現場では、できる限りの公平性、客観性が担保されなければならないということが要求されている。特に上級学校や就職先でこの公平性と客観性が担保されていないと、その学校からの進路保証が成り立たなくなってしまうだろう。しかし、学校現場では当たり前のように使用されている評定という用語が、教育学では評定がそれ独自のものとして扱われていないように思われる。

広辞苑によると、評定とは、一定の尺度に従って、価値・品などを定めること⁽²⁷⁾。

『現代授業研究大事典』によると、評価と違って指導との連環的サイクルとは離れたところで行われる対象についての客観的、普遍的な判断。評定は、評定者の主観や恣意を排除し、たとえ評定者が異なっても評定に変化のないことがめざされる。そのため、評定においてはこまかな評定尺度が考案され、できる限りの数量化が行われる⁽²⁸⁾。

上記の定義から評定が、客観化を担保したものだということがわかる。

(3) 評価と評定の違い

では、評価と評定の違いは何に求められるのだろうか⁽²⁹⁾。

ウィキペディアには、次のような記述がある。「評価は、いくつかの項目・観点に分けてなされることが多い。同じような意味合いで用いられる語として評定（ひょうてい）が挙げられるが、評定は「様々な評価を総合して、最終的に定めた値踏み」というニュアンスで、評価と評定は、厳密には別のものである。」⁽³⁰⁾

以上の定義付けからわかることは、評価と評定が教育学的に明確に区別されているわけではないということである。教育学では対象の関心をもつばら評価の方に向いていることである。一方、学校現場では、逆に評定の方に関心に向いていることが見て取れる。このことが、評価と評定の区別を厳然と区別してこなかった理由のひとつであろう。

(4) 評価と評定の関係

では、評価と評定の関係は、どのように捉えられているのだろうか。

『現代授業研究大事典』では、評価と評定の区別の意義を論じている。そして、次のように言う。「評価こそが教育実践の場で重視されなければならないことは明らかである。現実には評価と評定の区分は明確にされておらず、両者は混同されて用いられ、どちらかと言えば、評価の概念が評定の意味を強めて用いられている。評価は主観的であっても指導に役立てばよい。評価における客観性の重視は、教師の主体性の後退を意味する⁽³¹⁾。

『教育学小辞典』には、「評定」の項目はないが、評定尺度の項目がある。それによると、評価をより正しく行うためには、どうしても客観的に検査され測定された資料が必要であり、これを得る主要な方法が教育測定である。その測定の用具が各種のテストや評定尺度である。教育測定と教育評価は互いに一方の意味を含む同義に用いられることがあるが、同義的に用いるべきではない。歴史的に、主観的試験法を否定して生じた客観主義（教育測定）を弁証法的に発展させたのが教育評価なのだという⁽³²⁾。

要するに、教育評価が主観的なものに陥る危険性があり、客観性を担保するために、教育測定が必要とされ、その道具となっているのが評定尺度なのである。つまり、評定尺度とは、評価の客観性のための道具なのであり、評価が“主”で評定尺度が“従”の関係だと言えよう。

『[現代]教育学事典』にも「評定」の項目はないが、「評価とは何か」という項目の説明の中で次のように使用されている。「教育の評価は。子どもの学力や行動を評定する仕事であるとともに、このことをとおして教師の実践や教育条件を評定する仕事である。」⁽³³⁾

しかし、この説明では評価と評定の関係が明確ではない。続けて、「教育評価抜きの教育実践はあり得ない。教育実践が目標を持つ意図的、計画的行為であることに由来する。教師ならだれでも意識、無意識のうちにさまざまな場面で評価をしている。教育評価論は、意識、無意識のうちに行為されているこの仕事を自覚させる、教育実践の水準を向上させることだ」としている。したがって、発問だけではなく、テスト、通知表、内申書、指導要録、入試・進級制度などは、技術化され、制度化されている部分なのだという⁽³⁴⁾。

この説明からは、学校現場で使用されている（ウィキペディアの説明にあるような）実態としての「評定」というものは存在せず、元来の理念としての「評価」に内包されている。

そして、評価には評定尺度が必要だとするが、それは、教育実践が目標を意図的に追求

することから教育評価の必然性が生じ、評定尺度はその目標と表裏一体の関係にあるのだという。目標—評定尺度のあり方の違いによって評価の違いが生じてくるのだという⁽³⁵⁾。

この説明では、評価と評定の関係において評価が“従”で評定が“主”のように感じられるが、目標を定めた際に、ある評価を採用し、その妥当性のためにもそれぞれに見合った評定尺度が必要になるということなのであろう。評定尺度というのは、あくまで評価のための手段なのである。

ただ、この事典では〔評価と判定〕という項目で、学校現場で通用している「評定」の代わりに判定という用語が使われている。次のような説明である。評価行政：国家との関係において関わる教育評価のことでこの部分のことを評価行政といい、指導要録や入試制度などが重要部分となる。この部分が教育評価本来の目的である評価の仕事が子どもの励みの目標を与えたり、教育実践の良負を反省したり、教育条件の適不を占うという作用を失って、子どものできふできを選別する作用に変質してくる。これを評価と区別して判定という⁽³⁶⁾。

荘島宏二郎は、教育評価を測定 (measurement)、査定 (assessment)、評価 (evaluation) の関係から説き起こしている。能力 (学力でも良い) は、複雑で直接知覚することが難しい。それを査定によって、現象把握、現実理解、つまりいま何がどうなっているのか知ることである。その方法として、測定によって現象の諸側面を把握することであるのだという。測定とは、査定の1つ1つの行為でしかないため、現象の一側面しか把握できず、全体を捉えることが難しい。そこで多面的な測定結果を得て、現象を輻輳的に立体視することが重要になるのだという。測定の総体が査定となる。そして、その結果の意味を付与するのが評価である。また教育学のような客観的事象ではなくアーティファクトを扱う場合、その後の次のステップアップとしてどのようなアクションを起こすかといった評価を伴う判断の方が重要だという⁽³⁷⁾。

この場合では、査定が評定に近い用語で、測定が評定尺度に当たるものと言えよう。教育評価の理念を生かし、次に進めるために個々人の能力 (学力) を現象的に捉える必要があり、そのための個々の測定を行い、その総体としての査定するのだという。

さて、以上の考察から明らかになってきたことは、学校現場で使われている評定という用語が教育評価研究の世界では、独立して明確に位置づけられていないことである。評定尺度のように評価における手段であったり、「判定」という用語で代替されているように体制—反体制のイデオロギーの文脈に扱われていたり、何より、教育の理念からして教育評価に比して、扱われていなかったり、扱われていたとしてもマイナスの扱いを受けていたりする。ところが、学校現場では、評定が実態としてあるが、評定を含めて、全て評価として扱っているが、実質、それは評定であったりする。

教育評価研究における評価概念の理念と評定の軽視、学校現場における評価と評定の区別の曖昧さから教育評価論と学校現場との乖離が基底にあることが明らかになった。

5. 教育評価の変遷

ここでは、教育評価の成り立ちに焦点を当てて考察してみたい。

(1) 評価の類型

梶田毅一によると次のように分類している。①教育成果の評価であるのが総括的評価、②活動前の評価は事前的評価（子どもが現実にどのような発達の姿を示し、どのような能力や特性を現に持っているのかを見てとり、指導の前提としての一人ひとりの個性的なあり方を見てとるための評価である。）、診断的配置的評価（子どもの示す態度や発言、行動について、どの点はそのまま伸ばしていけばよいのか、どの点は特に指導して矯正すべきであるかを判断し、指導のストラテジー（方略）を立てる土台とする評価である。）、活動途上における評価を形成的評価（教育活動の中で子どもがどのように変容しつつあるか、を見てとり、一人ひとりに対する次の課題提示や指導のあり方を土台とする評価である。）、外部から活動の過程や成果や規定条件を吟味する外在的評価（教育活動自体がどの程度成功であったかを、子どもの姿自体の中から見てとる評価である。）としている⁽³⁸⁾。

ここで焦点としたい評価は、総括的評価と形成的評価であろう。総括的評価が結果を評価するのに比べて形成的評価は過程を評価するからである⁽³⁹⁾。

(2) 教育評価とその変遷

田中耕治によると教育評価とは、教育（学習）効果に対する判断行為一般をさす概念である。「試験」「考査」「測定」と区別する意図をもって、第2次世界大戦後に使用された用語だという。しかし、「試験」「考査」「測定」と混同・融合されていく経緯と重なっていく。つまり、評価の客観性を追求していく過程の中で「相対評価」に重きを置くようになっていた⁽⁴⁰⁾。しかし、優性思想や社会的ダーウィニズムに繋がるという批判もあり、1970年頃から再び教育評価の概念が復権してきた⁽⁴¹⁾。

その契機をつくったのがタイラーであった。タイラーは、試験の効果を教育的価値の実現の方向でコントロールしようとする問題意識が教育評価の意図的な使用へパラダイム転換がなされたという⁽⁴²⁾。その考え方を整理すると次のようになる。①評価の基準は教育目標、②教育目標は、高次の精神活動を含む目標群を含むべき、③教育目標は、生徒に期待される行動で記述すべき、④目標実現の度合いを知るための多様な評価方法が工夫されるべき、⑤目標未到達の子どもがいた場合には、治療的授業が実施されるべき、⑥以上のことは、カリキュラムや授業実践の改善につながる、⑦以上のことは、実践家と研究者の協力によって行われることが望ましい⁽⁴³⁾。

教育評価は、子どもたちを「ネブミ」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⁽⁴⁴⁾。従来は、教育測定と教育評価は厳然と区別されていて、教育評価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⁽⁴⁵⁾。要するに教育評価は、客観性・信頼性の担保ができなかったので、測定派が評価の方を取り込みながら（測定派の修正）、その方法の記述を行ってきたということなのである。

教育評価のこの客観性・信頼性を担保するため有力となったのが相対評価であった。学力分布が「正規分布」することが「ノーマル＝正常」ということになり、「正規分布曲線」になるようにテストが作成されるようになったのである⁽⁴⁶⁾。

こうした相対評価が教育評価を独占することによって、目標と評価を結びつけるという発想を阻害し続けていた⁽⁴⁷⁾ので、新しい評価観への転換が起こった。それが、ブルーム等の形成的評価やそのための道具としてクローズ・アップされた到達度テストの提唱などの影響を受けたものであった。梶田によると、それは、評価という視点から教育のありか

たを根本的に問い返し、評価という働きを有効適切に組み込んだ教育活動を、新たに構想し、設計し、実践していこうとするものである。学校、教師として持っているねがいやねらいをはっきりさせ、その観点から、従来の教育活動や指導のありかたを根本的に再吟味し、ねがいやねらいがよりよく実現するような活動や指導の計画を立てていく。ねらいやねがいを常に頭に置きながら子ども一人ひとりの現実の姿をさまざまなレベルで対話し、それに基づいて指導のあり方を軌道修正し、一人ひとりの子どもにも補充し、課題や深化課題を与え…といった着実な形での活動展開を図るものなのである⁽⁴⁸⁾。新しい評価は、それを導入することによって、教育方法のみならず教育を変革するものだというのだ。

(3) 日本の場合

日本の場合、戦前は絶対評価であったが、教師の主観が入り込む余地が大きく、恣意的なものとして批判され、戦後、アメリカの教育測定運動の影響を受けて相対評価が導入された。客観テストや正規分布曲線に基づく相対評価こそ客観的科学的評価法だともてはやされた⁽⁴⁹⁾が、周知の通りこの評価方法は、競争や選抜をあまり、対外的には上級学校への進学に使用されるようになった。

梶田によると、評価論、評価実践の展開史は、正規分布曲線の呪縛にどう対処し、どう乗り越えるかという基軸を中心に動いてきたという。昭和55（1980）、56（1981）年頃から指導要録の改訂によって到達度評価の原理に立つ「観点別学習状況」の欄が設けられ、ブルーム理論の紹介もあり、「指導と評価の一体化」へ進んでいったという⁽⁵⁰⁾。

そして、平成10年版（1998年）の学習指導要領にはいわゆる絶対評価を位置づけ、平成13年版（2001年）の指導要録の改善通知が出された。また、それ以前の平成元年版の学習指導要領で、新しい学習観、学力観が提示され、平成3年版の指導要録では観点別学習状況の評価が重視された。そして、平成13年版の指導要録において学習指導要領での各教科の目標の実施状況を観点ごとに評価するとされ、評価の観点を「関心・意欲・態度」、「思考・判断」、「技能・表現」、「知識・技能」に定められた。この観点別評価を指導に生かす評価として充実するよう、すなわち指導と評価の一体化が提言されたのである⁽⁵¹⁾。

こうした教育評価の考え方は、第2次世界大戦後、アメリカの進歩主義の考え方とともに紹介されたのが最初だという⁽⁵²⁾。1969年の「通信簿論争」を経て、梶田一たちによって、ブルーム学派の教育評価論が紹介され⁽⁵³⁾、「相対評価」から「到達度評価」の成立に至ったのだという。

6. 観点別評価の学校現場への導入に関する問題

観点別評価のうち、最も問題だといわれている「関心、意欲、態度」を取り上げて考えてみよう。高木展郎も次のように述べる。「しかし、この「関心、意欲、態度」も評価の方法が難しく、それまでのペーパーテストによる評価が一般化された中で、この「関心、意欲、態度」をどのように評価すべきかが、学校教育において大きな問題となった。そこには、評価の客観性という、評価の公正性と公平性が大きく関わっている⁽⁵⁴⁾。しかし、問題点を指摘するだけで、どう対応すればよいのかは述べられていない。暗にペーパーテストに慣らされている教員の意識の問題だとも読めるが、その後の客観性、公正性と公平

性が担保できないから必要最低限としてのペーパーテストに頼っているということがわかっていないように思われる。

また評価と評定の区別もついていない。そもそも評価研究者によると、梶田は、その批判に対して、タイラーの言に沿って、測定ではなく評価が大切なのであり、これは、客観的で厳密であることを第1の条件とするのではないと主張する⁽⁵⁵⁾。評価は何よりもまず客観的で厳密でなくてはならない、という思い込みがまだ根強い⁽⁵⁶⁾と逆批判を行い、評価の重要性を強調する。田中は、例えば形成的評価について、授業改善のためであって、成績付けには使われないと述べる⁽⁵⁷⁾。ここに問題の根っ子があると言えよう。前記、3、4で論じたように、評価研究では、評価の重要性を強調するが、学校現場において厳然と存在する評定については、軽視するか、もしくはまるで問題外なのである。

「関心、意欲、態度」の評価について、挙手の数を評価に入れていることが批判されてきた。文科省も公的に批判してきた。象徴的なのは、次のような批判である。「その評価については、情意面にかかわる観点であることなどから、目標に準拠した評価であることが十分理解されていなかったり、授業中の挙手や発言の回数といった表面的な状況のみで評価されるなど、必ずしも適切とは言えない面も見られる。」⁽⁵⁸⁾

梶田は、「関心・意欲・態度」の評価の批判に対して、タイラーの主張を援用して「測定ではなく評価が大切」だと述べ⁽⁵⁹⁾、結局、「関心、意欲、態度」を目指して明確化することを主張する。そのためにブルームらが開発した「教育目標のタキソノミー（情意的領域）」を援用すべきという⁽⁶⁰⁾。また、「構え」の形成が学校教育を通じて追及されなくてはならないと唱える⁽⁶¹⁾。測定ではなく、評価が大切だといっても、ブルームなどのタキソノミーを参考にしても、「構え」が大事だと唱えても評価が曖昧で、評定重視の学校現場では単なる精神論としか聞こえないであろう。

田中は次のように述べる。「指導要録における「観点別学習状況」欄の「関心・意欲・態度」項目をめぐるのは、教育現場でさまざまな対応が現れている。その最も特徴的な点：「関心・意欲・態度」とは「学習態度」（授業への参加態度、発達態度、学習への真面目さなど）のことと理解して、たとえば授業中での子どもたちの「挙手」のあり方が着目されることもある。挙手率、挙手の姿までが評価のチェックポイントになる。「関心・意欲・態度」を「態度主義」と批判するといっても必ずしもこの問題への具体的な解決方法を打ち出すには至っていない。この問題を解くために、認知と情意の関係を認識の深化にとって必要不可欠な契機として両「要素」の関係を具体的に理解することである。」⁽⁶²⁾

田中の主張は、現場教員の教育評価の明確な理解が解決に繋がると考えているように見える。認知と情意の相関関係、すなわち内発的動機が学力を向上させるという論理なのだろうが、では、その内発的動機を教員が評価してしまったらそれは外発的動機に成り代わるのではないだろうか。また理解ができれば、この問題を解けるとは思われぬ。現場の教師にしてみれば、それをどう客観的に評価するのかということなのである。情意と認知の相互関係など経験値でわかっている。情意を客観的に評価しなければならないから困っているのである。客観性を担保するために仕方なく行っているだけなのだ。

こうした批判の理由は、この評価への現場の認識不足だということなのであろう。だが、現場の感覚でいえば、その評価研究で述べられる程度の評価は（すべての教員とは言わないが）、日常的に、口頭等で行っているのである。問題なのは、そのような評価を指導要

録等の文章で明確化し、結果として、その評価が評定に繋がってしまうことなのである。そうならば、客観性を担保せざるを得ない。特に中学校の場合、内申書が個々の生徒の進路に多大な影響を及ぼす。「客観性」、「公正性」、「公平性」が著しく問われる。さらにその説明責任を果たさなければならない。そこで、不本意ながら挙手の回数やレポートの数などを取り入れざるを得なかったということなのである。それとも現場の教員、全てがペーパーテストでの知識・理解のみで生徒たちを評価、評定するだけだと思っているのであろうか。それこそ、現場教員に対する侮辱以外のなにものではないであろう。つまり、上記、3で論じた形成的評価のような評価の客観性をどう担保するかという課題が横たわっているのである。

次に、この評価と評定を一体化してしまっているにもかかわらず、観点別評価を行うことの問題点をいくつか挙げてみよう⁽⁶³⁾。

第1に、結果の評価ではなく過程の評価（形成的評価）であるため「客観性」、「公正性」、「公平性」、が著しく阻害されることである。例えば、観点別評価において一番問題があると言われている「関心、意欲、態度」の評価を行う場合、多人数でその場で動いている生徒たちを一瞬にして評価しなければならないこと。特に中学校の場合、教科担任制なので、数多くの学級を受け持つ。1時間目の3年A組の授業でのA君の活動での「関心・意欲・態度」の評価（→評定）と6時間目の3年B組でのB君の活動での「関心・意欲・態度」の評価（→評定）に差が出たとき、どのように本人及び保護者に説得できる説明責任を果たせるのだろうか。1時間目と6時間目の授業を同じように行ったという保証ができるのだろうか。それで「客観性」、「公正性」、「公平性」が担保できるのか。では、同じ時間の同じ学級内では可能なのかと言え、例えば、1時間目の3年A組で、C君とDさんの「関心・意欲・態度」を評価（→評定）してDさんの方が高い評価（→評定）が出たとする。Dさんの方が教師から見て積極的に学習に取り組んでいるように見えたといったときに、B君は内向的なので、教師から見えるような形では「関心・意欲・態度」の評価（→評定）が見られないのだと本人及び保護者から抗議を受けたら、本人及び保護者を納得できる説明責任が教師にできるものであろうか⁽⁶⁴⁾。単に口頭で「Dさんは〇〇のことに関心と意欲を持って取り組んだが、B君の場合は、今一步だった」というような説明ではどうして説得されないであろう。そこで「客観性」、「公正性」、「公平性」を担保するために数値で表すということになる⁽⁶⁵⁾。実際問題として、評価（→評定）なので数値で表さざるを得ないのだ。数値で表すとした場合、その「関心・意欲・態度」の達成度なるものをどう判断し、評価（→評定）のカットポイントをどこに据えるのかという問題が持ち上がる⁽⁶⁶⁾。精密なルーブリック評価表を作成したとしても授業中に生徒たちの一クラスで考えた場合でも最大40人の「関心・意欲・態度」を見てとり、「客観性」、「公正性」、「公平性」を担保した評価（→評定）を行い、さらに本人及び保護者に説明責任を果たせるだけのことができるのかは疑わしい。また、中学校、高等学校の場合でいえば、英語なり数学なりを一人の先生がひとつの学年すべて受け持つことがない場合が多い。その場合、3年A組のE君の「関心・意欲・態度」をG先生が評価（→評定）し、3年B組のFさんの「関心・意欲・態度」をH先生が評価（→評定）した場合、教師間の整合性がとれるのだろうか。精緻なルーブリック評価表を作成し、綿密な打ち合わせをしたりしたとしても相手の先生が評価（→評定）した生徒を実際に見ているわけではないのだ。これで「客観性」、

「公正性」、「公平性」を担保できるのだろうか。さらに、受け持ったクラスが週に1回の場合、1学期間の間にそれほど授業があるわけではない。その少ない授業での「客観性」、「公正性」、「公平性」な評価ができるものなのだろうか⁽⁶⁷⁾。結局、挙手した回数とかレポートの数、枚数とか後の限られた“証拠物件”となるもので評価(→評定)せざるをえないのだ。

第2に、教師は“審判や判定者”ではなく、“プレイヤー”であることが忘れられている。

教師が授業を行っているとき、それは生徒と同じプレイヤーなのだ。教える、段取りをする、指示をするといった教師のプレー中に冷静に「客観性」、「公正性」、「公平性」をもって、生徒たちを評価(→評定)することが可能なのだろうか。毎回の授業に教師とは別に評価(→評定)する者が必要であろう⁽⁶⁸⁾。そのため、教師の仕事の主たるところが教えることではなく、評価することとなっている。教師は、「評価」することではなく、「教える」ことが本来の教育行為だと感じている。教科の力をつけさせたい。「教科の立場」につくとき、「関心・意欲・態度」を重視する評価に否定的になるのだという⁽⁶⁹⁾。

第3に、教師の負担が大きいことである。評価に関する仕事が以前に比べて3倍ぐらいに、いや、何倍にもなったという。学校のブラック化の要因のひとつであろう。筆者の知り合いの小学校、中学校の教員の学期末の通知表づくりで、忙しさが増したという。学校のブラック化は、部活だけの問題ではなく、この観点別評価の導入も大きいと考えている。それでいて、「説明責任」のために自己目的化し、生徒のためになっているのか疑問だという。「労が多くて功が少ない仕事はやりたくない」という意見も出されたという⁽⁷⁰⁾。教師の働き方改革が叫ばれている中で、観点別評価の問題を指摘する意見は少ない。

第4に、本来、教えるべき教材の内容の準備より評価基準に時間が取られてしまうことである。日本の教師は、特に授業以外の業務(生活指導、公務分掌等)に従事していて、(評価研究者のように)評価だけに専念できるわけではない。東京都のある小学校の教師は、次のように述べたという。「教師の主観でどうにでもつけられる性質を有する観点別評価を厳密にやろうとしたら恐ろしく大変でとても不可能なので主観的判断でつけているのが小学校の現状だ」という⁽⁷¹⁾。中学校では入試があるため、この不可能と思われる状況に取り組みざるを得ない⁽⁷²⁾。真面目に取り組みれば取り組むほど“無間地獄”に陥らざるを得ない。

第5に、第4にも関わることではあるが、評価(→評定)作業に多くの時間と労力が割かれるために、教材研究に割かれる時間と労力が削がれてしまうことである。結果的に、教師の教える能力の軽視に繋がってしまうことになる。教師は、「評価」することではなく、「教える」ことが本来の教育行為だと感じている。教科の力をつけさせたい。「教科の立場」につくとき、例えば、「関心・意欲・態度」を重視する評価に否定的になる⁽⁷³⁾。

第6に、授業内容が評価に規定されてしまうことである。授業の内容よりも評価(→評定)に重点が置かれることによって、評価(→評定)しやすくするために授業内容が薄っぺらなものになってしまうことである。50分の授業の中に、浅い問い、授業内容の流れの中で意味のない作業、話し合いやプレゼン活動を盛り込む結果、あたかも“多様な関連性のない具が入った丼ぶり”のようになり、いったいこの授業では何を生徒に身に付けさせたかったのかわからなくなる。目的と手段の転倒である⁽⁷⁴⁾。

第7に、教科の性格になじまない評価基準もあるのではないのかという疑問である。観

点別評価は、全教科に渡って同じ「関心・意欲・態度」、「思考・判断・表現」、「資料活用技能」、「知識・理解」という評価（→評定）を行うように定めている。例えば、自然科学と違い、価値基準が多様な社会科に当てはめることができるのか。ましてや中等教育以上に適用できるのだろうか。各教科に内在した評価基準ではなく、頂点にこの4観点の基準があり、それに各教科が整合性を保つよう作成されている。それも各教科の単元毎に作成されるということは、かなり無理な要求だと言わざるを得ない⁽⁷⁵⁾。このように、教科に内在した評価基準ではなく、教育心理学系で定説となった評価基準を一律に全ての教科で、それも小学校、中学校、高等学校に当てはめるという事態に対して、恐らく作成した人たちは、全ての教科のエキスパートなのだろう⁽⁷⁶⁾。また例えば、数学的思考・判断と社会科的思考・判断は同じものなのか。数学的思考・判断ができると社会科的思考・判断ができるのか。社会科でも歴史的思考と公民的思考は同等のものなのか。考えれば考えるほど疑問が湧き起ってくるのである。

さらに、生徒側から見た観点別評価の問題点について述べることにする。

第1に、生徒の内面に当たる観点別評価を評価できるのかという問題がある。例えば、「関心・意欲・態度」は、生徒の内面管理に繋がらないのだろうか。生徒からすれば、評価（→）のために常に教師の前で、「関心・意欲・態度」があるように見せなければならぬ⁽⁷⁷⁾。本来、生徒がその教科や単元に「関心・意欲・態度」が持ち得なくてもあるように見せかける場合は、それを教師が見抜けるのだろうか。よしんば、見抜けたとしても客観的な根拠なしにそれを評価（→評定）に反映できるのだろうか。実際、観点別評価対策として、塾や予備校が生徒のための対策まで行っている。以下は、「関心・意欲・態度」におけるその事例である。

どの教科も、観点1は、**関心・意欲・態度**です。数学なら数学、理科なら理科という教科に、**・関心を持っているか・意欲を持っているか・どういった態度で臨んでいるか**ということの評価します。意欲があるかどうかや、関心を持っているか、こんなことは、ペーパーテストでは、評価できません。**テストの点が良ければ、・関心を持っている？・意欲がある？**確かに、そうかもしれませんが、テストの点が悪い人は、関心がない、意欲がないとは、言い切れません。ですから、この観点だけは、誰でも、Aを取ることができます。Aを取るためには、意欲があること、関心を持っていることをアピールしましょう。**・授業中に発表をする。・積極的に質問をする。・宿題以外の勉強にも取り組む。**こんなことが、アピールになるのではないのでしょうか？しかし、ここで勘違いする人がいます。良いですか？**・授業中に騒ぐ、話を聞かない。・宿題をしない、提出物を出さない。**こんな人は、意欲があると言えますか？関心を持っていると言えますか？たまに、授業中に紙飛行機を飛ばしているくせに、上のようなアピールをしてくる人がいます。ハッキリ言って、逆効果です。変なアピールはしない方がまだマシです。**授業と宿題が1番大切**であることを肝に銘じておきましょう⁽⁷⁸⁾。

評価する側は生徒のどのようなところをみて、「関心・意欲・態度」が十分にあると判断しているのでしょうか？チェックされるポイントのいくつかの例を教科ごとにご紹介します。授業態度、提出物（ノート・ワーク・小テスト）、忘れ物、発言など。授業態度、

課題への取り組み (ファイル・ワーク提出)。授業態度、持ち物、提出物、定期考査、ノートへの書き込みなど。授業態度、服装、持ち物、定期考査など。ほとんどの教科で「授業態度」というのは必ずチェックされています。他には宿題などの提出物、小テストの結果や忘れ物などのようです。また国語や数学などの主要5教科では、「関心・意欲・態度」の観点には定期考査の評価は含まれず、反対に実技4教科では定期考査は「関心・意欲・態度」の観点で評価される傾向があるようです。ここまでで、「関心・意欲・態度」の観点はどのような基準で、どこを見られて評価されるかをご紹介します。では、どのようにして評価をあげるか。そのためには「一生懸命やっています」「やる気があります」ということをアピールするにつきます。提出物の期限内での提出は当たり前、忘れ物も0、授業も積極的に参加して授業に関する質問、発言を増やす。これらさえ、しっかり続けていけば、自信のない教科でも必ず良い評価に変わります。反対に、これらが徹底してできていないと、高い評価は期待できません。提出物などは、「期限を忘れないようにメモをしておく」や「後回しにせず、早めにすませておく」などの工夫や意識改善を促してあげることが第一歩です。特に「関心・意欲・態度」は積み重ねがものをいう評価観点。小さなことから、コツコツと取り組んでいきましょう。

授業態度は、主に授業中に授業を受ける姿勢・気持ちを評価します。観点でいうと「関心・意欲・態度」の評価に大きく影響するのも授業態度でしょう。「授業を受けている・聞いている」だけでは少し足りません。「授業内容に“関心が強く”“意欲的に”受けている・聞いている」と思われて、ようやくいい評価がもらえるのです。反対に「関心も低く、意欲もない」と評価されれば、低評価は免れません。具体的には、以下の行動が「関心が低く、意欲がない」と判断されます。授業開始時に席に着いていない。授業の道具 (筆記用具、教科書、ノート、プリント) が机に出ていない。ボーッとしている / 授業とは関係ないことを考えている / よそ見。居眠り / あくび。猫背 / 膝を立てる / 頬杖 (ほおづえ) をつく。立ち歩く / 消しゴムをちぎって投げる。携帯電話をいじる。チョコやアメなど、お菓子を食べる。友達との授業に関係ないおしゃべり / 友達と手紙交換。先生を呼び捨て、「センコー」呼ばわり、あだ名呼び。注意に対する無意味かつ無根拠の言い訳、反論。

上記の他にも、たくさんあるのは言うまでもありません。授業中の態度は普段の家庭での生活が影響する場合があります。授業中にボーッとしている、居眠りをしてしまうなどは睡眠時間が少ないことが原因のケースもありますし、姿勢や態度そのものは、家庭での生活がそのまま反映することもありますから。以上が授業態度に関する「内申点を下げる、やってはいけない行動」です。今回ご紹介したのが全てではありません。紹介された中で、「うちの子が当てはまるのはなかったわ」と思われた親御さんも、安心するのはまだ早いのです。「下げる行動をしていない」=「内申点があがる」ではありませんから。ここでポイントになるのは、その授業に対しての「やる気」の部分です。そのため授業態度で評価されることが多いのです。

【行動例】 忘れ物は絶対にしない。授業開始時に席に着いている (着いていなければ遅刻です)。机には授業に必要な道具 (教科書、ノート、プリント類など) を出しておく。

姿勢を正す（前を向く、背筋を伸ばす、ほおづえをつかない）。私語はしない。宿題やテスト前の提出物は期限内の提出が基本中の基本。期限内に提出したからといって成績があがるわけではありません。反対に期限を過ぎてしまうと即マイナス評価につながりますので、注意が必要です。「1つでも提出漏れや期限遅れがあったらAはつかない」という意識をお子さんと共有しておきましょう。また、先生が「出しても出さなくてもいいよ」と言った場合、「出さなくてもマイナスはしないが、出せばプラス」というニュアンスのようです。当然出すべきですね。提出期限を2段階に設定し、早く出したらプラス評価するという場合もあるようですので、このチャンスは是非とも活用したいもの。こうしたことも、授業中に口頭で伝えられるわけですから、それも含めて授業への取り組みは大切です。お子さんの「聴いていなかった」「知りませんでした」は全く言い訳にはなりません。なお、「関心・意欲・態度」では、「提出物を出したかどうか」を重く見るようです。出来が良いかどうかは、むしろ「技能」「知識・理解」「思考・判断・表現」での評価対象となる場合も多いようです⁽⁷⁹⁾。

第2に、生徒の側にとって、「思考・判断」も思考も何をもってして思考といえるのかの定義が多様であり、学習指導要領に定められた思考以外の思考が認められた場合、それは思考として評価されるのかという疑問が残り、また他の観点別評価も含めて調査書に合わせた思考を行っているように示すことになる。結局、調査書を意識して、自分が即できること、つまり授業態度をよくしたり提出物を出したりするようになる。難しい数学の問題にチャレンジすることではなく、教師が喜ぶようなノートや態度をつくることに専念する危険がある。成績の良い子ども自分を伸ばすことにエネルギーを注がないということになってしまっている⁽⁸⁰⁾。

7. 観点別評価の学校業幾への導入に関する根本的な疑問

田中は、形成的評価は、授業改善のためであって、成績付けには使われないという⁽⁸¹⁾。しかし、現実問題として、評定と結びつかない評価にどれだけ有効性が保たれるのだろうか。「あてにならない評価」⁽⁸²⁾とされたり、また評価と評定が真逆であったら説明責任を求められるだろう。さらに評定と別の評価をつけるほど教員は暇でもない。結局、評定を想定した評価を行うしかなくなる。実際、評価は評定となり、調査書に記載されている。また、観点別評価の評定への転換の方法は、既に各学校や教育委員会レベルで行われている。例えば、伊藤敏雄「なるほど！中学校の通知表・成績表の評価の付け方」⁽⁸³⁾によると、次のような説明を行っている。「通知表の中の社会科の成績なら「関心」「思考」「表現」「知識」の4つの観点について、「絶対評価」で評価されています。観点別評価は、「関心」のような授業中の挙手の回数や調べ学習での取り組みから、「知識」のようなテストの点数まで、それぞれA～Cの3段階で成績をつけます。そして、これらの組み合わせで次のように通知表の評定が決まっています（観点は教科によって異なります）。**通知表の評定の決まり方の例**：4つの観点の評価の組み合わせ（順不同）→通知表の評定・AAAA、AAAB→「5」・AAAC AABB→「4」・ABBB、BBBB、BBBC→「3」・BBCC、ACCC→「2」・BCCC、CCCC→「1」※同じ評価の組み合わせでも、学校や教科によ

て評定は異なります。このように通知表の評定が決まるため、テストの点数が90点(＝A評価)で良かったとしても、関心などのその他の観点が普通(＝B評価)だった場合、4つの観点の評価の組み合わせは「BBBA」となり、通知表の評定は「5」ではなく「3」になってしまうのです。つまり、単純にテストの点数だけで通知表の評定が決まっているわけではないのです。」

文科省からも、観点別評価とは別に評定を行うことは引き続き必要であるのだという。結局、評価の結果を評定に「転換」という。すると田中は、この「観点別」評価から「評定」に転換するためのルールを開発・普及することが、今後求められるというのである⁽⁸⁴⁾。しかしこれは、観点別評価ではなく「観点別評定」なのではないだろうか。形成的評価ではなく「形成的評定」なのではないだろうか。ではなぜ、教育評価が評定に「転換」されてしまうのだろうか。その理由は、いくつか考えられる。

第1に、日本の学校教育が置かれている社会的状況を考慮に入れていないことである。日本での教育評価論の言説は、あたかも実験室で行われた実験のように、もしくは思弁的な“あるべき論”としての教育評価論なのである。1960年代以降の高度経済成長期以降、日本の産業構造が変化して日本の学校教育を取り巻く環境が変化した。良い悪いは別にして、学校教育を経て自分の人生を豊かに実現するというコースが“王道”となり、学校教育における良い評定への要求は強まっていった⁽⁸⁵⁾。そうした社会状況下において、進路保障に繋がらない教育評価が成り立つものなのだろうか。しかし、例えば中内敏夫は、受験教育を入試制度改革で解決しようという案に対して、次のように述べている。「従来のような意味での優等生や受験学力を結果しないような教育評価のあり方と機構をつくりだすことを目標にしなければならぬであろう。」⁽⁸⁶⁾また田中は、次のように主張している。「教育評価として立論するとは、評価基準の転換を教育の内容事項と外的事項の両面にわたって実現することによって、教育評価の社会的文脈に働く論理をも組み替えることを要求することである」⁽⁸⁷⁾。要するに、教育評価を導入することによって、教育方法や学校教育とそれを取り巻く環境も変えられるということである。例えば、入学試験制度について、田中は、次のように述べている。「[「相対評価」に代わって、「目標に準拠した評価」に転換していく中で、入学試験制度も「資格型」が模索されている。そうすると、入学試験制度とは上級学校と下級学校を分断する装置ではなく、「接続」する装置として、その資格内容を満たしているのかを判定し、以後の指導にも役立てられるという発想が生まれる。」⁽⁸⁸⁾しかし筆者の見解では、逆に社会状況や社会的文脈に依存して学校教育が成り立っているのである。したがって、学校教育の中でのみ通用する教育評価にそのような力があるとは思えない。教育評価は、その時代の社会状況や文脈を表す一面でしかない。評価のあり方を「選抜型」から「資格型」に転換するといっても、では、全国の大学を資格型にできるのだろうか。例えば、東京大学をはじめとした全ての大学をフラットにして、大学資格を満たせば大学卒ということで社会が、企業の採用で納得するだろうか。例えば、京都大学が、資格要件を満たした者を全員入れられるのだろうか。2倍も3倍も入学者が増えた場合、教員の数も含めて大学の設置基準の法律も変えねばならないだろう。大学のキャパシティを越えた場合はどうするのだろうか。資格要件を難しくしたならば、結局、それは“選抜”でしかない。

第2に、観点別評価は、形成的評価の流れの中で定められたものである。学びの結果で

はなく学びの形成過程を評価する。そのためであろうか、各教科の観点別評価表には、「～している」という表現が見受けられる。例えば、社会科において次のような表現である⁽⁸⁹⁾。

「日本史B」の内容のまとめり（学習指導要領の大項目）「(4) 近代日本の形成と世界」における評価規準に盛り込むべき事項である。（「学習評価参考資料」より一部抜粋）

(4) 近代日本の形成と世界」における評価規準に盛り込むべき事項

関心・意欲・態度

近代国家の形成と社会や文化の特色に対する関心と課題意識を高め、意欲的に追究している。

思考・判断・表現

近代国家の形成と社会や文化の特色から課題を見だし、国際環境と関連付けて多面的・多角的に考察するとともに、国際社会の変化を踏まえ公正に判断して、その過程や結果を適切に表現している。

資料活用の技能

近代国家の形成と社会や文化の特色に関する諸資料を収集し、有用な情報を適切に選択して、読み取ったり図表などにまとめたりしている。

知識・理解

近代国家の形成と社会や文化の特色についての基本的な事柄を、国際環境と関連付けて総合的に理解し、その知識を身に付けている。

上記の下線部の部分は、ing を表現しているのであろうが、この評価（→評定）を行ったときには、過去である was でしかない。この奇妙な表現は、観点別評価の表現として定着してしまっているが、未だに違和感がぬぐえない。例えば、関心・意欲・態度の追求しているのは、今も継続して追求しているということなのか、そうだとしたらこの評価を職員室で行っているのにそれがなぜわかるのか。それでなければ、ある時期の事なのか、ではどの時期のことなのか。ある生徒はこの評価を行った時期には追求していたかもしれないが、この評価を行った直後に追求を怠った場合は、どうするのか。別の生徒が追及していないと評価をした後に追求していったらどうなるのか、その時期を明記しないのに公正性は保たれるのか。そもそものこの問題の根底の原因は、この教育評価の理論は、静止している対象を捉えるものであるにもかかわらず、動いている対象を捉えようとするためである。この理論は、動画を捉えることができるのだろうか。生徒のある学びの形成過程の一断面をスナップ写真のように捉えるに過ぎないのではないだろうか。そのため、動画に近づけるために何枚、いや何十枚ものスナップ写真を撮り続けなければならない⁽⁹⁰⁾。

第3に、観点別評価相互の関連性を評価（→評定）する手立てがしっかりと示されていないことである。例えば、長谷川康男は、観点別評価に基づいてペーパーテストを作成する困難さを主張している。「どれが、「関心・意欲・態度」で、どれが「思考・判断」なのか。どれが「思考・判断」で、どれが「観察・資料活用の技能・表現」なのかが分からなくなる。迷う」⁽⁹¹⁾。生徒は、総体として学びを形成しているにもかかわらず、それを要

素に分けて評価（→評定）しなければならない。各観点の関連性も生徒の学びの形成過程を捉えるものなのかという疑問符が付く。そこで、教員が“恣意的に”各観点別評価（→評定）を“振り分ける”ということになってしまうのである。そこに教員としての疑問を持ちながら、多くの労力を割くことになってしまっているのである。

8. おわりに—グローバルな評価基準に抗して—

ここまで批判的な論を展開してきたが、筆者は、観点別評価で述べられている内容それ自体を否定しているわけではない。それを「客観的」な評価にし、それを基準にして評価（→評定）にしてしまうことを批判してきたのである。学校教育は、指導と評価が一体化していなければならないというが、実際の学校は、そのように評価が前面に押し出されているものではない。評価は、評定と明確に分け、評定を必要最小限のものとし、評価は、教員が生徒を様々な観点から見眼を養うものとして、口頭で生徒に伝えていけばよいのだと考える。

ところが、この観点別評価は、今後、さらにグローバルな背景をもってグレードアップした形で導入されようとしている。1997年に OECD が着手した PISA 調査において、次のように報告されている。「人生における生徒の成功はいっそう広い範囲のコンピテンシーに左右されることがわかってきたという。そして、OECD の DeSeCo プロジェクトによって、キー・コンピテンシーが定められた。それは、次の3つに代表される。

①相互作用的に道具を用いる。②自律的に活動する。③異質な集団で交流する。」⁽⁹²⁾

その枠組みの中で、国立教育政策研究所は、「21世紀に求められる資質・能力」としてまとめた⁽⁹³⁾。その報告書によると、「資質」とは、子供が学んでいくために持っている潜在的な力と考える。「能力」とは、「資質」を使って実際に学ぶことで、子供自らが育て、自覚的に活用できるようになった力と考えてみることもできる⁽⁹⁴⁾。

それを学習に当てはまると次のように言えるという。①学び初めには学習に使う手段、学び終わりでは学習内容も含みこんだ次の学習のための手段である。したがって、方法知でありつつ内容知を含みこんだもの。②知識の質向上のために必要不可欠な手段かつ目標となる。「手段」は、知識の質を上げるために資質・能力を使うことが必要不可欠であり、「目標」質の上がった知識やそれらを統合したものの方・考え方、知識を仲間とともに作り替えられるという態度等を含みこんだ資質・能力が目標となること。③「資質」を中心に人格（価値・態度等）に関わるものである。なお、この際、価値を教えて子供の「資質・能力」に組み込むか、あるいは、価値は学ぶ対象としておいて、その受容は子供の判断に任せるかは重要な検討課題であるとのこと。知識だけではなく、スキル、更に態度を含んだ人間の能力であるとのこと⁽⁹⁵⁾。続いて、松下佳代の次のような論を引用している。「新しい能力」は、認知的要素（知識、スキル）だけではなく、人格のより深部の非認知的要素（動機、特性、自己概念、態度、価値観など）を含もうとしている。能力は、領域固有か普遍（汎用的）を水平軸（広さ）で捉える。非認知的能力は、「市民性」「持続可能な社会づくり」等の価値に関わる教育は、垂直軸（深さ）で捉える⁽⁹⁶⁾。そして、裁量の拡大で多様な実践が生まれることを推奨しつつ、評価基準の共有で一定の質保証を行うとしている⁽⁹⁷⁾。さらに、教育課程の基準に、子供を主語として「(子供が)～をできるよう

になる」という形で教育目標を明記するというのである⁽⁹⁸⁾。つまり、ここでは、一人一人の子供が知識やスキルを使って、学んで獲得した（できるようになったこと）やその学んでいる態度、さらにグローバルな基準で価値を持つ活動に主体的に参画することまでもが評価（→評定）の対象となるということなのだ。こうした、ある時代、ある地域の価値あるとみなされる（ここでは、欧米的価値になっているであろう）ことまでもが評価（→評定）されてしまうという恐ろしさを抜きにしても、どのようにこれら进行评估（→評定）するのであろうか。

福田誠治は、こうしたグローバルな潮流に対して、次のように主張している。①西欧近代に登場した科学的知性の限界を認め、評価する側の能力にかかわらず素人でもわかるというような「評価の可視化」を中止すること。人間評価には、評価する側の教養と能力が大きく関わってくることを認め、判断とその説明は専門家である現場の教師に任せるべきで、現場にいない教育行政関係者や保護者の介入は極力避けるべきである。②評価の観点の多様化。教科の得点とか総合点という考えを止めること。数値にした途端に、意味の世界はこの世から遠くに雲散霧消してしまう。③現在の評価システムでおろそかになっている個人の人間らしさなどの議論を活発化させること。グローバル人材など企業に役に立つ、金儲けになることだけが評価ではない⁽⁹⁹⁾。

けだし、卓見であろう。

(注)

- (1) 拙論「学力問題を考える－高校教師の視点から－」（埼玉学園大学紀要 経営学部編 第16号 2016年12月）p.114.
- (2) 「[「ペーパーテスト偏重」脱却なるか 高校でも観点別評価]」（日本教育新聞 2018年11月26日）、「高校に観点別評価導入 中教審部会決定 知識偏重の授業から脱却へ」（日本経済新聞 2019年1月28日）
- (3) 「学習評価，3観点に整理 中教審WG報告 学習の計画性も」（日本教育新聞 2019年1月7日）
- (4) 梶田叡一『教育評価〔第2版〕』（有斐閣，1992年）p.85.
- (5) 梶田，前掲書，pp.86～87.
- (6) 田中耕治『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，2010年）p.12.
- (7) 梶田叡一『形成的な評価のために』（明治図書，2016年）pp.132～133.
- (8) 牧野茂『巨人軍かく勝てり V 9の秘密』（文春文庫，1983年）
- (9) 牧野，前掲書，pp.30～31.
- (10) 牧野，前掲書，pp.32～33.
- (11) 梶田，前掲，形成的評価のために，pp.89～90. 附属高校では，生徒やカキュラム上の環境が揃っており，形成的評価の研究校として取り組めば，成果が上がったように見えるだろう。
- (12) 梶田，前掲，形成的評価のために，p.91.

- (13) 梶田, 前掲, 形成的評価のために, pp.112~113.
- (14) 梶田, 前掲, 形成的評価のために, p.115. 数字は筆者。
- (15) 梶田, 前掲, 形成的評価のために, p.118.
- (16) 梶田, 前掲, 形成的評価のために, pp.116~119.
- (17) 梶田, 前掲, 形成的評価のために, p.99.
- (18) 梶田, 前掲, 形成的評価のために, p.110.
- (19) 『広辞苑 第7版』(岩波書店, 2018年)
- (20) 山田栄, 唐沢富太郎, 伊藤和衛, 大浦猛編『教育学小辞典』(協同出版, 1974年)
- (21) 吉本均責任編集『現代授業研究大事典』(明治図書, 1987年)
- (22) 梶田, 前掲, 教育評価, p.1.
- (23) 田中耕治『教育評価』(岩波書店, 2008年) p.15.
- (24) 田中耕治『新しい「評価のあり方」を拓くー「目標に準拠した評価」のこれまでとこれからー』(日本標準, 2010年) p.10.
- (25) 田中, 前掲, 教育評価, p.15.
- (26) 各種の教育学の辞典や事典を調べたが, 評定尺度という用語は掲載されているが, 評定そのものはほとんど掲載されていなかった。評定ではなく, 測定になるのであろう。
- (27) 前掲, 『広辞苑 第7版』
- (28) 前掲, 『現代授業研究大事典』
- (29) 『現代授業研究大事典』では, 英語で評価を evaluation, 評定を valuation と表記しているが, 手元にある研究社の『新英和中辞典』によると, evaluation の意味は, 査定, 評価で, valuation の意味は, ① a 評価, 値踏み, b 見積もり [査定] 価格, ② (人物・才能) などの評価, 品定め, 判断となっている。どちらにも評価の意味があり, 中辞典では, そのニュアンスを受け取ることはできないが, 本稿の目的が学校現場への導入であることを考えると, 現場に導入する際のこの曖昧さは, 様々な誤解を生じる恐れがあるだろう。英英辞典やもっと専門的なニュアンスを調査するというを学校現場で行えるとは思えないからである。
- (30) 「評価」ウィキペディア (Wikipedia) (2018年現在)。なお, ウィキペディアには「評定」の項目もあり, 次のような記述になっている。「学校における評定は, 3段階・5段階を用いることが多い。教師は学習の様子, 宿題, 定期考査の結果などを資料として児童・生徒を評価し, 評価した結果を総合して評定を決定する。公立学校においては, このときの評価資料を評価する方法として2000年ごろから絶対評価が用いられている。評定は学期ごとに通知表に記載されて児童・生徒やその保護者に通知される他, 学年末に指導要録に記録され, 保存される。また受験の際の調査書にも記載され, 内申点として用いられる。」この内容は, 学校現場における評定の意味に近い。
- (31) 前掲, 『現代授業研究大事典』
- (32) 前掲, 『教育学小辞典』
- (33) 青木一, 大槻健, 小川利夫, 柿沼肇, 斎藤浩史, 鈴木秀一, 山住正己編『[現代]教育学事典』(労働旬報社, 1988年)

- (34) 前掲, 『[現代] 教育学事典』
- (35) 前掲, 『[現代] 教育学事典』
- (36) 前掲, 『[現代] 教育学事典』
- (37) 荘島宏二郎「古典的テスト理論—科学的対象としてのテスト得点—」(植野真臣 荘島宏二郎『学習評価の新潮流』(朝倉書店, 2010年) pp.49~51.
- (38) 梶田, 前掲, 教育評価, pp.1~4.
- (39) 読売巨人軍の査定と類似する。
- (40) 相対評価の源流は, アメリカの「メジャメント」が源流。統計的な尺度で客観性を求めていった結果であった(田中, 前掲, 教育評価, p.21.)
- (41) 田中, 前掲, 教育評価, p.15.
- (42) 田中, 前掲, 教育評価, p.24.
- (43) 田中, 前掲, 教育評価, pp.26~27.
- (44) 田中, 前掲, 教育評価, p.34.
- (45) 田中, 前掲, 教育評価, p.35.
- (46) 田中, 前掲, 教育評価, pp.42~43.
- (47) 田中, 前掲, 教育評価, p.97.
- (48) 梶田, 前掲, 形成的評価, pp.38~39.
- (49) 梶田, 前掲, 教育評価, p.121.
- (50) 梶田, 前掲, 教育評価, pp.316~318.
- (51) 高木展郎「指導要録の改訂と学習評価の変遷」www.mext.go.jp/b_menu/shingi/chousa/shotou/111/shiryo/_icsFiles/afiefieldfile/2015/08/19/1360907_8.pdf
- (52) 田中, 前掲, 教育評価, p.33.
- (53) 田中, 前掲, 教育評価, p.216.
- (54) 高木, 前掲。
- (55) 梶田, 前掲, 教育評価, p.177.
- (56) 梶田, 前掲, 形成的評価のために, p.110.
- (57) 田中, 前掲, 教育評価, p.123.
- (58) 教育課程審議会答申(2000年12月4日)
- (59) 梶田, 前掲, 教育評価, p.177.
- (60) 梶田, 前掲, 教育評価, pp.179~p.181.
- (61) 梶田, 前掲, 教育評価, p.118.
- (62) 田中, 前掲, 教育評価, p.114.
- (63) 以降, 評価と評定が一体化してしまっている状況を「評価(→評定)」と示すことにする。
- (64) 恐らく, 評価研究の専門家であればできるのであろう…。
- (65) A中学校の沢木の証言「02年は, 5教科の評価がすごく大変だった」。「日常的活動を点数化しないと評価は出せないと思う。」理由として, 生徒や保護者に対する「説明責任」に言及している(金子真理子「中学校における評価行為の変容と帰結—教育改革の実施過程に関する社会学的研究—」(『教育社会学研究』(第72集, 2003) p.112, 115.)。

- (66) 金子, 前掲論文, p.114.
- (67) 金子, 前掲論文, p.117.
- (68) 前記, 2でも示した通り, 読売巨人軍では, 監督一人が査定していたわけではない。数人のコーチがいて, 監督の手助けをしている。監督もプレイヤーなのであるから。また, 某中学校の教師から聞いた話ではあるが, その中学校の体育教師は, 生徒たちの動きを評価(→評定)するために観察に徹し, 指導をしなかったそうである。指導しなかったのではなく指導できなかったのであろう。
- (69) 金子, 前掲論文, pp.118~119.
- (70) 金子, 前掲論文, p.117.
- (71) 金子, 前掲論文, p.116.
- (72) 金子, 前掲論文, pp.116~117.
- (73) 金子, 前掲論文, pp.118~119.
- (74) 教育評価を変えることによって, 授業変革ができるのだという主張もある。例えば, 「パフォーマンス課題を設定して授業を計画するようになってからは「教科書を暗記させないと, 力はつかない」という呪縛から解き放たれた」という実践もある(三藤あさみ・西岡かなえ『パフォーマンス評価にどう取り組むか—中学校社会科のカリキュラムと授業づくり—』(日本標準, 2010年) p.15.)。そういう場合もあるだろう。しかし, 例えば, 社会科教育実践において教材研究の内容からやクロスカリキュラムやテーマ学習など多彩な授業方法論の採用から授業を変える場合の方が多くであろう。筆者の感覚から言えば, 自身の授業について, 評価を変えるまで授業を変えるという自覚がなかったということに関して, それまでの授業実践に対しての疑義を持たざるをえないのである。
- (75) 実際に筆者は, 観点別評価が導入された直後の教育実習生の指導案を見せてもらったことがあるが, 単なる割り算の計算練習にも観点別評価の基準を作るのに悪戦苦闘した跡が見られた。割り算の計算練習にどのようにしたら「関心・意欲・態度」の評価基準が作成できるのだろうか…。
- (76) この根底的な問題には, 教育学のヒエラルヒーも影響を与えていると考えられる。各教科教育学の上に「教育学」が君臨しているという構図である。それだからこそ, 各教科があまり疑問を抱かず, 観点別評価の評価基準を受け入れて, 各教科の評価を何とかそれに合わせて作成しようと四苦八苦することになる。しかし, 例えば社会科教育に関して言えば, 戦後から数えてもすでに70年の歴史がある。相当な蓄積を積んできているはずである。自らの教科から独自の評価を提出しても良い時期だと考えている。
- (77) 極端なことを言えば, その光景は, 映像で見る朝鮮民主主義共和国の朝鮮労働党大会における金正恩の前での党幹部の態度を思い起こしてしまうのである。
- (78) ScienceLab 富田塾のホームページより。
- (79) 以上, school post のホームページより(石井知哉執筆から)。その他, 「技能」や「思考・判断」の対策も参照してほしい。観点における学力を向上させることよりも, いかに観点別評価の調査書の点数を上げる対策にそのエネルギーが費やされているかがわかるであろう。

- (80) 金子, 前掲論文, p.119.
- (81) 田中, 前掲, 教育評価 p.123.
- (82) 金子, 前掲論文, pp.120~121.
- (83) all about <https://allabout.co.jp/gm/gc/66170/>
- (84) 田中, 前掲, 新しい「教育評価のあり方」を拓く, pp.59~60.
- (85) さしあたり, 堀尾輝久『日本の教育』(東京大学出版会, 1994年) IV 経済と教育を参照。
- (86) 中内敏夫『増補 学力と評価の理論』(国土社, 1976年) p.130.
- (87) 田中, 前掲, 教育評価, p.169.
- (88) 田中, 前掲, 新しい「教育評価のあり方」を拓く, p.21.
- (89) 神奈川県教育委員会『学習評価のてびき』(平成25年1月) より。下線部は筆者。
- (90) そもそも教育評価の理論自体が欧米のものであり, 西洋哲学から発展したものであることは論を待たない。ところで, ハイデガーは, 西洋哲学の存在観について次のように論じている。「在る, 存在するというこの意味は, [アリストテレスにとって] 元来, 制作されてあるということに存するからである。」ただ, この存在の意味は, 付帯的なものであったが, 「存在論的な探究がさらに発展してゆく過程の中で, 実在性や現実性という漠とした不特定の平均的意義へと墮してゆき, そういったかたちでやがて認識論的な問題構制の基本的な視点となるに至る。続いてまたこの視点に基づいて初めて, 対象を「自然」と理論的に規定する場合の「客観性」が生じるに至り, これが問題の中心をなす存在意味となってゆくのである。」(ハイデガー『アリストテレスの現象学的解釈『存在と時間』への道』(平凡社, 2008年) pp.98~100.)。つまり, ここではピュシュ(自然)や運動という現象について古代ギリシャがもっていた根源的な経験は, パルメニデスにおいて狭められ, 不変不動のものを「見る」ことが存在物とのかかわりの中で特に優位を占め, それが以後の西洋の存在観を規定するところとなったのである(ハイデガー, 前掲書, p.192。(高田珠樹『『ナトルプ報告』の成立とその位置』)。木田元によれば, 〈自然〉はもはや生きて生成(し消滅)する自然であることをやめ, 超自然的な原理にのっとった制作の無機的な〈質料〉になってしまうのだ(木田元編著『『存在と時間』の構築』(岩波現代文庫, 2000年, p.211.)。そもそも, この存在観に規定された西洋の認識論が静止している対象しか捉えられないものなのではないだろうか。それなのに, 元々“静止している”理論や基準で生徒たちという, “動いている対象”を評価せよと論じているのである。ここに大きなギャップがあるのではないだろうか。それ故にであろうか, 梶田は, 次のように述べる。「評価が「格づけ」や「選別・差別」になってしまうのは, 子どもの姿を確定的なものとしてとらえるからである。…子どもの姿を暫定的なものとして, つまり, 手の打ち方次第でいかようにも変わっていくものととらえる。…子どもの姿そのものも, またそのある面を測定したり評価したりした結果も, 「色」であって「色」ではない。「空」であって「空」ではない」(梶田, 前掲, 形成的な評価のために, pp. 187~188.)。いくら, 現場の教師に対して生徒を固定的に見るなど精神論での啓蒙を繰り返してきても固定的にしか評価しない教師に対して, 最終的には, (無意識にも) 西洋の枠組みから脱した仏教の枠組

みから言及しているのである。ここに梶田は、自らの西洋の論理の限界を見てとっているのかもしれないが…。実際には、学校現場が、生徒を固定的に捉えるのは、そもそも観点別評価が形成的評価と言われながら固定した静止した論(枠組み)からきているため、そのように評価してしまうのであり、そこを学校現場に対して“精神論”で何とかしろと言っているにすぎず、無意識的にも仏教の枠組みを発したのかもしれない。

- (91) 長谷川康男『評価を生かした問題解決学習でめきめきつく確かな学力』(学時出版, 2004年) p.30.
- (92) ライチェン, サルガニク編著『キー・コンピテンシー 国際標準の学力をめざして』(明石書店, 2006年) pp.200~203.
- (93) 国立教育政策研究所編『資質・能力 理論編』(東洋館出版社, 2016年)。この報告書を執筆したメンバーのほとんどが教育心理学系の研究者であり、参考文献も学校教育の大部分を占める教科教育の文献がほぼなく、教育心理学, とりわけ認知心理学系の文献が多いのが気付きである。
- (94) 国立教育政策研究所, 前掲, p.39.
- (95) 国立教育政策研究所, 前掲書, pp.67~68.
- (96) 国立教育政策研究所, 前掲書, p.32.
- (97) 国立教育政策研究所, 前掲書, p.29.
- (98) 国立教育政策研究所, 前掲書, p.30.
- (99) 福田誠治『ネオリベラル期教育の思想と構造 書き換えられた教育の原理』(東信堂, 2017年) pp.508~509.

Received : April, 2, 2019

Accepted : June, 12, 2019