

教育におけるアカウンタビリティと 新しい評価の動向

奈須 正裕

1 学習評価とカリキュラム評価

まず、「教育評価、それは可能か？」という特集タイトルにおどろいた。そんなもの、可能に決まっているではないか。もし不可能なら、我々はどのようにして学生の単位認定を行っているというのだ。

こう言われると、今度は本誌編集担当者の方がおどろくだろう。「単位認定？いやそういう話じゃなくてですねえ。」そう、編集担当者がここでイメージしているのは、個々の学生の学習状況の評価ではなく、特定の学部なり全カリ、あるいは立教大学の教育というものはどれほどの教育効果を上げているのか、それを明らかにしたいが果たして可能か、という意味での教育評価なのである。

というわけで、まずは整理しておこう。一口に教育評価というが、そこには当初私がイメージした学習評価と、編集担当者がイメージしたであろうカリキュラム評価という少なくとも2つのレベルがある。そして、本特集でもっぱら問題となるのは、カリキュラム

評価の方なのである。

これらは同じ評価という言葉を使っているが、評価を受ける対象が正反対である。すなわち、学習評価は特定のカリキュラムの下で学んだ個々の学習者の学習状況の適否であるのに対し、後者はそのような学習状況をもたらしたカリキュラムの適否である。

もともと、個々の学習者の学習状況が明らかになってこそ、カリキュラムの適否や問題点、改善点も明らかとなるのであり、両者は照応関係にある。

たとえば、1988年の教育改革以降、イギリスの公立学校では、7、11、14歳の各時点で、全国統一の学力テストが実施されている。その結果は、個人の学習状況の評価となると同時に、一種のカリキュラム評価として学校ごとにも集計され、その学校が成し遂げた教育効果の証として一般に公表される。

この情報が社会的に大きな意味を持つのは、イギリスの公立学校が学校選択制を取っているからだ。子どもや保護者は学校選択に際し、当然のことながらこのカリキュラム評価に関する情

報を重視する。さらにシビアなことに、イギリスでは入学者の人数が学校予算の配分に反映される。カリキュラム評価の結果は、学校の命運を左右する一大問題なのだ。

我が国でも近年、学校選択の考え方やその具現化の方途としての学区の自由化や弾力化が徐々にではあるが広がりを見せており、各学校がどのような成果を生みだしているかに関心が寄せられ始めている。これを受け、市区町村単位で統一の学力テストを実施し、学校ごとの結果を公表しようとの動きも出始めている。もっとも、目下のところ我が国では競争が比較的ゆるやかなものであること、競争の結果を予算その他と連動させないことなどの点において、イギリスなどとはかなり状況が異なる。

2 アカウンタビリティ

このような動向は、1970年代にアメリカで生まれたアカウンタビリティ運動に端を発する。アカウンタビリティという言葉は、もともとは会計学の用語で、資金が違法に費やされたり使い込まれることのないよう保証することを意味した。これが教育の世界に持ち込まれ、税金を払っている市民は、公教育がその支出に見合うだけの教育効果をあげているかどうかを知る権利を有し、教育関係者は教育効果を報告する責任があるとする考え方が生まれてきた。このアカウンタビリティの評価

モデルは、少なくとも以下の3つの段階を含む。

1) 目標アカウンタビリティ

第1は、カリキュラムの目標それ自体の評価である。すなわち、設定された教育目標が妥当であるかに関する評価であり、目標アカウンタビリティと呼ばれる。この段階の評価に対し責任を負うのは、教育目標の策定責任者、我が国の公立小中学校でいうならば、文部科学省であり、文部科学大臣である。

近年、いわゆる「ゆとり教育」の是非をめぐる議論が盛んであり、一部の数学者を中心に学習指導要領の即時停止を求める署名運動までなされたが、これなどは文部科学省が提示した学校教育の目標なり、そこで設定された学力の志向性や水準が不適切である、すなわち目標がアカウンタブルではないとの判断に基づく行動と理解できよう。

ここで重要なのは、我が国の公立小中学校では、学習指導要領によって教育目標なり目指すべき学力が詳細かつ明示的に表現されており、その課程を修了した際に身に付いているであろう事柄、その間に生じるであろう子どもたちの変化をかなり具体的に想定することができる点である。それが、目標なり学力がアカウンタブルであるかどうかを、誰もが共通の足場に立って議論することを可能としている。

これに対し、大学教育の目標、より具体的には特定の学部なり全カリ、あるいは立教大学の教育目標とは具体的にどのように陳述されているのか。もちろん、建学の精神や学部設立の趣旨としては述べられており、大学案内等のメディアによっても広く公開されている。しかし、ではその課程を修了した暁には、学生にどのような変化が生じることを保証しようとしているのか。少なくとも、それを具体的に評価しうるほどにまで詳細に記述できているかとなると、まだまだあいまいさが残ってように思えるのである。

2) プログラム・アカウンタビリティ

第2は、目標に基づいて実際の教育実践が計画・実施されているかどうかの評価であり、プログラム・アカウンタビリティと呼ばれる。この段階の評価に対し責任を負うのは、プログラムの開発・選択責任者、我が国の公立小中学校でいうならば、学校長である。

たとえば、授業内容の偏向といった表現で、学習指導要領の趣旨から大きく逸脱した授業実践が批判され、是正を求めるといったことがなされてきたが、そこでは現場で開発、選択された教育プログラムが、目標に照らしてアカウンタブルかどうかということが主要な論点となるのである。

大学教育に目を転ずると、特定の学部なり全カリのカリキュラムの詳細、あるいはそれを構成する個々の講義の

内容は、教育目標とのどのような関連において、あるいは教育目標のどの部分の達成を担うものとして開発・選択され、そこに位置付けられているのか。また、それは十分にアカウンタブルなものと言えるか、が問われる。たしかに、こういった事柄はカリキュラムに関する議論として、実に熱心に、また慎重になされてきたと言えるだろう。しかし、残念なことに、目標にあいまいさが残る以上、その実現を目指すプログラムのアカウンタビリティについても、評価に一定の困難さがつきまとうのは必然と言わざるを得ないのではないだろうか。

3) 成果のアカウンタビリティ

第3は、プログラムによって無理なく達成できるであろうだけの成果を、現実に個々の学習者の内に実現し得たかどうかの評価であり、成果のアカウンタビリティと呼ばれる。この段階の評価に対し責任を負うのは、当然個々の教員である。

先に述べた、学習評価との照応関係をもつカリキュラム評価とは、まさにこの段階に相当する。このことからわかるように、カリキュラム評価とは、個々の学習者の学習状況の評価に先立ち、そもそも目標それ自体が妥当であるか、また開発・選択されたプログラム自体は目標に対し適切といえるか、といったことが問われるのであり、そこを度外視にして学習者の状況だけを

云々するのでは不十分なのである。

大学教育で考えるならば、個々の講義における学習内容の学生への定着の度合いが問題とされるわけであるが、ここでやっかいなのが、大学教員に与えられた講義実施への自由度の高さである。同じ講義名の講義でさえ、そこで指導される内容や運営の方針、さらには目指される達成度が異なることは少なくない。したがって、ある特定の講義なり教員がもたらした成果のアカウントビリティの評価も、何とも難しいものとなる。では、テキストや教材、さらにはテストまでも共通化し、指導の内容はもとより方法までも制御する方がよい教育となるかという、それもまた微妙である。少なくとも、個々の教員の専門性が存分に生かされない。そのことから、かえって講義の質の低下が危ぶまれるとの批判の出ることは、容易に想像のつくところであろう。

3 スタンダード準拠評価

つまるところ、教育目標が具体的に明示的なものとならない限りにおいて、いかなる評価も不可能ないしは不徹底である。大学教育に限らず、教育目標というどうしても抽象的なものとなりやすく、カリキュラムの開発やその評価といった実際の作業の拠り所とならない、お題目的なものとなるが多かった。すでに述べてきたことからわかる通り、評価論は目標論

の裏返し、それも学習者の姿の変化をとらえるという、何とも具体的な行為を伴う裏返しであり、実際化という側面を持つ。

この点について、具体的で明示的なものとする動きの強すぎることは、具体的で明示的にとらえやすいもの、典型的にはペーパーテストによって測定しやすい部分、数量的に見積もりやすい側面へと、教育の目標や学力論、さらにはカリキュラムや授業実践を矮小化したり歪曲する危険性があるのではないか。多少のあいまいさは、むしろ教育の豊かさにとって必要なものではないか、との指摘のあることは、十分に予測し得るところである。

実際、アカウントビリティ運動の隆盛に伴い、学校や教師がペーパーテストでの得点向上に血道を上げ、深い思考や個性的な探究を求める授業展開への敬遠、さらには芸術を始めとして、ペーパーテストにのりにくい教科や領域への軽視が生じ、結果的に教育がバランスを欠き、子どもの十全な発達保障を危うくするということもあった。近年の我が国における「学力低下」論が、結果的にドリルや受験学力に偏った学力論に世論を誘導する危険性ははらんでいるのも、同様の現象と理解できよう。

あるいは、大学教育において目指す学力というのは、初等中等教育段階などとは違い、要素的に、あるいは単純に記述できるようなものではなく、も

つと高次で複雑なものである。したがって、具体的で明示的な評価などにはおよそのらないのだ、との批判もあるかもしれない。

しかし、そういった学力論の重視はもはや大学教育に限らない。近年、初等中等教育段階における学力論は、世界的に見ても、三桁の足し算が速く正確にできることや、単語や漢字を書けることだけではなく、問題解決力や思考力、判断力、表現力など、いわゆる高次の技能を重視する方向へと進んできている。我が国の「ゆとり教育」も、少なくとも本来はそのような趣旨のものであった。

しかし、そういった高次の技能は、○×や多肢選択で答えられるような単純なテストでは十分に評価できない。では、どうすればいいのか。この難問を解決すべく、1990年代以降、オーストラリアやアメリカ、イギリスを中心に、スタンダード準拠評価ないしはルーブリックと呼ばれる、新しい評価の考え方と道具立てが生み出されてきた。

スタンダード準拠評価ないしはルーブリックとは、求める能力なり技能の水準の特徴を一般化して示した言語表現と、それを充足する学習者の学習活動における具体的な現れ（作品や事例）の具体例との組み合わせによって、評価基準を示す方法である。以下に、理科で用いられている評価基準の一例を示す。

・評価基準：レベル3（西オーストラリア州の理科の評価基準）

生徒は微粒子構造を用いることにより、材料の特性を説明するモデルや概念を理解していることを示すことができる。

・生徒の事例：生徒と教師の会話の記録（水と砂糖の混合について）

教師：何が起こっていますか。

生徒：砂糖が水に入って消えてしまっています。

教師：そのとおりです。これをどう説明したらよいですか。

生徒：えーと、水の中では粒子がぎっしりと詰まっているのだけど、…砂糖はそれらの隙間にちょうど入り込むことができると思う。砂糖は水の中では、より小さな粒子となって、水の中に入っていくのでしょうか。

もちろん、ここに示した会話は例であって、実際にはさらに多様な姿が立ち現れていい。重要なことは、この例示があることにより、実際に教室で出会うであろう多様な現れについても、かなりの客観性を持って評価を行えることである。この工夫により、一般的であり、その意味であいまいさがつきまとう高次の技能の評価が一層確かな

ものとなった。現在、我が国も含め、スタンダード準拠評価ないしはルーブリックによる評価は、伝統的な客観テストによる評価を補完し、あるいはそれと置き換わる勢いで現場に普及し始めようとしている。従来より、高次な技能を重視し、それがゆえに評価が困難とされてきた大学教育においても、このようなアプローチは大いに可能性があるのではないだろうか。

なす まさひろ

(本学文学部教育学科教授)