Dr Negishi's Assessment Clinic
評価クリニック

観点ごとの総括をどうするか

根岸 雅史 Negishi Masashi
（東京外国語大学教授）

1. 増えた評価資料

観点別・絶対評価というシステムが導入されてから，英語の教師は従来のペーパー・テストだけでなく，様々な評価方法を採用するようになった。その代表的な例が，授業中の観察やインタビュー，宿題などであろう。この結果，評価対象の特性に応じて多様な評価方法の中から最適なものを選択するようになってきている。

2. 観点ごとの総括

2-1. テスト得点によらない方法

では，このように集められた観点別学習状況の評価結果は，どのように総括されるのであろうか。これまでは，多くの場合，定期試験の点数を合計し，それを相対的に解釈すればよかったが，観点別・絶対評価においては，集められた複数の評価結果をもとに観点ごとの ABC を決定しなければならない。

まず，現状として，どのような方式がとられているかをみてみよう。平田 (2003) によれば，評価基準を用いて，1 回 1 回の評価活動において，「おおむね満足できる状況（B）」に達していれば○，達していなければ×と判断し，その数をベースに総括する方法がある（本稿ではこれを【頻度法】と呼ぶ）。例えば，ひとつの評価規準を 3 回適用した場合を考えると，すべて○ならば A，○が 2 つまたは 1 つならば B，すべて×ならば C という具合である。

この方法に対して，ABC を数値化して総括する方法がある（これを【数値換算法】と呼ぶ）。国立教育政策研究所（2002）によれば，「…評価規準に照らして行った評価結果の状況を数値によって表せるようにし，この数値から総括を行う事例も見られた。例えば，A を 3，B を 2，C を 1 とする場合で，評価規準に照らして行った評価結果が『A，A，B』という状況は平均点 2.67（小数点第 3 位を四捨五入），『A，A，B，B』という状況は平均点 2.5 という数値で表されることとなる」というわけだ。

2-2. テスト得点による方法

ところが，【頻度法】や【数値換算法】が適用されているのは，主に「コミュニケーションへの関心・意欲・態度」の評価に対してである。おそらく，これは「コミュニケーションへの関心・意欲・態度」の評価では，「評価基準を用いて，1 回 1 回の評価活動を評価する」ことになるからであろう。

これ以外の評価の場合は，定期試験などのテストが用いられるために，別の方法が用いられることが多いようだ。それは，【テスト得点合算法】というような方法である。この方法では，基本的に，それぞれの観点の評価をテストにより行い，その得点を合計して，ある基準により，ABC を判定するというものである。例えば，「理解の能力」でいえば，「聞くこと」および「読むこと」に関するテストの得点を合計して，それが 80％以上であれば A，50％以上 80％未満であれば B，50％未満であれば C という具合である。もちろんこの線引きは，各学校により異なっている。

3. 観点ごとの総括方法の「評価」

一見すると，昨年度より中学校に導入された絶対評価は，設定された目標に到達しているかどうかを見る「到達度評価（criterion-referenced grading）」のようである。この「到達度評価」型の絶対評価では，基準（もとじゅん）と呼ばれる最低基準を設定

し，これに達する学習者を「到達者（master）」，それ未満の学習者を「非到達者（non-master）」と区別することになる。したがって，ここには本来2つのカテゴリーしか存在しない。ところが，昨年度から導入された絶対評価をよく見てみると，奇妙な点に気がつくのである。それは，この絶対評価では，到達者と非到達者という二分法ではなく，ABCという「三分法」がとられているという点である（Aは「十分満足できる」状況，Bは「おおむね満足できる」状況，Cは「努力を要する」状況）。このABCを通常の到達度評価に当てはめると，AとBが到達者，Cが非到達者ということになる。

これを走り高跳びにたとえるならば，到達者と非到達者という二分法では，評価基準はオリンピック標準記録のようなものである。これを越えていれば，どれだけ越えていようが，ぎりぎりであろうが，参加資格があるということになる。そして，これを越えていなければ，どれだけぎりぎりでバーを落としても，まったく跳べずにバーを落としても，ともに参加資格はないということになる。ところが，今中学で行われている絶対評価では，単に跳べているだけでなく，どれだけ余裕を持って跳べているかも示されなければならない。

こうした視点から，観点ごとの総括方法を「評価」してみる。まず，【頻度法】であるが，これを走り高跳びとしてみると，基準となるバーを確実に何度も越えているとAになるというものである。しかしながら，何度もバーを越えられるものが，そのバーのはるかに上を越えている保証はないのである。少なくとも，これは一般の人が抱くAのイメージとは異なるのではないか。

次に，【数値換算法】である。これは，個人の記録を平均しているために，ここから得られる数値は，能力の違いを比較的よく反映しているように思われる。ただし，実際に得られる数値は1～3の間の連続データであるために，例えば2.5という平均値がAかBかは，主観的かつ恣意的な判断となる。

では，もうひとつの【テスト得点合算法】はどうであろうか。これは基本的な考えは，【数値換算法】と同様であるが，もとになっているデータがABCでなく，定期試験などのテスト・データである点が異なっている。この方法の問題点としては，【数値換算法】と同様のABCの線引きの問題のほかに，ABCの判断がテストの難易度に影響を受けているという点がある。これは，他の2つの総括方法と大きく異なる問題である。他の2つの総括方法では，まず始めに能力の判断を○×やABCで行っている。しかし，【テスト得点合算法】では，各問題の難易度がABCの定義に必ずしも対応しているわけではない。そのために，ABCの得点の線引きが事前に決まっていれば，ABCにより規定される能力はテストの難易度によって変わってしまう。たとえば，BとCの線引きが50％にあるときに，テスト方法としてTF法を使うことがある。ところが，TF法では，鉛筆を転がしても50％の正答確率があるわけで，これでは，Bは何の能力も保証していないことになる。

線引きの基準が学校全体として決められているというような場合は，引かれた線の上下に質的な差があるようにテストが設計されていなければならない。ところが，テストの難易度をあらかじめ予測して作成することは容易ではないのだ。これに対して，テストを実施した後に，ABCの線引きを行う方法もある。つまり，このテストでは，AとBの線引きは90％，BとCの線引きは60％というように判断するということである。また，このアプローチでは，教師の指導経験に基づき「ぎりぎり到達者」を何人かと「ぎりぎり非到達者」を何人かあらかじめ決めておき，彼らの出来をもとに線引きの数値をはじき出すという方法も提唱されている。

実際には，「日常点」による【数値換算法】と【テスト得点合算法】が合わさった【ハイブリッド型】も存在していて，実態はさらに複雑かも知れない。そもそも，どの方式を採用するかは，何を評価するかにもよるであろう。いずれにしても，それぞれの特性をよく理解して，総括法を選択し，運用する必要がある。

【参考文献】

○国立教育政策研究所教育課程研究センター『評価規準の作成，評価方法の工夫改善のための参考資料（中学校）―評価規準，評価方法等の研究開発（報告）―』国立教育政策研究所，2002年
○平田和人『中学校英語科絶対評価の方法と実際』明治図書，2003年