



気になるテスト作り

根岸 雅史 Negishi Masashi
(東京外国語大学)

観点別・絶対評価が中学の英語教育に導入されてからしばらくは、絶対評価とは何か、それぞれの観点では何を見るのかといったことが多くの英語教師の関心事であったような気がする。ところが、今年に入って、その関心は、観点別・絶対評価での「テスト」の在り方に移ってきているようだ。評価の枠組みは変わったのに、テストは変わらない——この矛盾に英語教師が気づき始めたのかもしれない。

そこで、今回は、観点別・絶対評価の枠組みにおいて用いられている英語のテストが、どんな問題点を抱えているのか、そのいくつかに焦点を当て、考察する。

1. 設計図のないテスト

まず、「もの作り」の一般的な手順を考えてみよう。作ろうとするもののコンセプト形成から始まり、それを具現化するための設計図の作成を行う。そして、場合によっては、それをもとに試作品が作られ、最終的なプロダクトに至る。家作りにしても、車作りにしても、こうしたプロセスをたどり、ものが作られていく。いい「もの作り」のために、とりわけ重要なのが設計図の作成である。家の設計図もなしに、窓を作り始めたりすることはないし、車の設計図もなしに、ドアを作り始めたりすることもない。

ところが、これがテスト作りとなると「設計図」の作成を行わない人たちがあまりにも多い。テスト作りでは、多くの人が、まず問題を書くことから始めているのである。どんな全体像のテストとなるのかもわからずに、個々のテスト項目を作成している。こうして作られるテストは「英語力」をただ漠然と測っているだけであり、観点別・絶対評価においては意味がない。

言語テストの場合は、全体図を描いた「設計図」のことを test design と呼び、そのパーツの詳細についての取り決めを書いたものを test specifications という。ここにはおよそ以下のような事柄が含まれることになる。

- ・評価計画に沿ったテストの役割についての記述
- ・そのテストで測ろうとする能力や知識のリスト
- ・それぞれの能力や知識をどのような方法で測るかについての記述
- ・リーディングやリスニングでは、どのようなテキストを用いるかについての記述
- ・それぞれの小問に必要な小問の数
- ・それぞれの小問の重みづけ（配点）
- ・それぞれの小問の重みづけ（配点）
- ・採点基準

定期試験などでは、少なくともこのようなことを決めずに、実際の問題作成に入るべきではないと考える。これが proficiency test ともなれば、これ以外にも、使える単語や文法項目・言語機能などのリストも確定しておく必要がある。

このような「設計図」を作ることで、テスト作成者は、自分の指導の中で何が重要で、それをどの程度まで身につけていることを生徒に期待しているのかを考えることになる。明確な「設計図」があれば、何の能力（または、知識）を測ろうとしているのかわからないような問題はなくなるはずである。

2. 「総合問題」の意味

観点別・絶対評価にあっては、ある到達目標に照らして、個人がその目標に到達しているかどうかを見ることになっている。したがって、何の能力（または、知識）を測っているのかわからないようなテ

スト問題は、観点別・絶対評価にあつては、上に述べたとおり、利用のしようがない。この意味で、観点別・絶対評価になってまずその存在意義がなくなるはずであったのが、いわゆる「総合問題」であるが、この「総合問題」は定期試験からいまだに消え去ってはいない。ご存じの通り、「総合問題」では、あるまとまった英語の文章が提示され、その中にある単語の発音を問うたり、単語の意味を問うたり、文法事項を問うたり、文章の内容を問うたりしている。このように、「総合問題」にはばらばらのテスト・ポイントが含まれているために、その結果の意味を解釈することは難しく、その得点は、観点別・絶対評価の枠組みでは行き場を失うのである。

往々にして、この「総合問題」の得点は、「理解の能力」として扱われていることがあるが、その中身を見れば、「理解の能力」を測っているわけではないということは明らかである。とりわけ定期試験における「総合問題」では、文章の内容理解を問う問題の比率は意外と少なく（まったく含まれていないこともある）、用いている文章は「総合問題」を「理解の能力」のテストとするための見せかけにすぎない。

3. 文法問題

観点別・絶対評価においては、どういった能力や知識を測っているかが明確になっていなければならないのは、いわゆる「文法問題」でも同じである。一般に、「文法問題」は、「総合問題」とは異なり、1つの大問の中に様々なタイプの問題が含まれているようには見えない。それは、「文法問題」は、それぞれの小問における問題形式がたいして統一されているからである（例えば、「空所補充問題」「並べ換え問題」「適語選択問題」という具合である）。しかし、それが実はくせ者である。「問題形式」が統一されているからといって、「テスト・ポイント」が統一されているわけではないのである。例えば、すべてのテスト項目が適語選択問題であっても、動詞の時制を問う問題が入っていたり、前置詞を選ぶ問題が入っていたり、単語や熟語の知識を問う問題が入っていたりすれば、この大問の結果から、生徒が何ができて何ができていないのかは見え

てこない。また、同じ並べ換え問題でも、あるところでは句の内部構造を問うているのに、あるところでは内部構造は問題にしていないといったようなことがある。となれば、こうした大問の出来から、生徒の文法知識を的確に診断することはできないのである。さらに、この種の文法問題から導き出される診断コメントは、たとえば「空所を補充する能力がある」とか「正しい語を選ぶ能力がある」などとなってしまい、およそ英語の指導目標からはかけ離れてしまう。これでは、このテストに適切な診断機能を期待することはできないのである。「文法問題」では、「問題形式」をそろえるだけでなく、それぞれの小問で、文法のどのような知識を測りたいのかを明確に意識する必要がある。

4. 厄介な「線引き」

観点別・絶対評価における評定の出し方については、本誌 Vol. 4 において取り上げたとおりである。相対評価では、合計点を出して、それを上から5段階の比率にあわせて割り振ってあげれば、評定が出た。これに対して、観点別・絶対評価では、それぞれの観点ごとに、基準に達しているかについて判断しなければならない。しかしながら、その基準の決定は恣意的である。例えば、基準を5割とするのか、6割とするのかなどは、教師自身が決めたり、学校で統一の基準があったりする。観点別・絶対評価にテストを用いた場合には、この基準に合うようにテストが作られていなければならない。実際、基準の「線引き」の作業は、テストの難易度とも絡んで、かなり厄介な作業である。たとえ同じ実力・同じ「線引き」であっても、テストが易しければ基準に到達しやすくなるし、テストが難しければ基準に到達しにくくなる。また、TFで理解を問うような問題であれば、「鉛筆を転がしても」5割は正解することになっており、これをもって基準に達しているとは判断すべきではない。いくらいいテストを作っても、この「線引き」のさじ加減1つで、結果は大きく異なってきてしまう。今までの惰性に流されることなく、一度立ち止まって自分の「線引き」の妥当性を検討してみたいかだろうか。