

アメリカ・イリノイ州の 教育評価テストについて

研究開発部試験方法研究部門 平 直 樹

平成8年3月から平成9年1月までの約10か月間、文部省長期在外研究員として海外出張を命ぜられ、その期間の大半の時期をアメリカ合衆国イリノイ州のイリノイ大学アーバナ＝シャンペーン校で過ごした。本稿は、イリノイ州におけるIGAP (Illinois Goal Assessment Program) と呼ばれる教育評価テストに関連して見聞した、アメリカのテスト事情についての報告である。

1 アメリカの教育改革とテスト事情

よく知られていることではあるが、最初に、教育とテストに関する社会的な背景について、我が国と異なるアメリカ合衆国に特有な事情について簡単に触れる。

1-1 教育改革の背景

1960, 70年代を通じて、カレッジ・ボード (CB, CEEB) が教育テストサービス (ETS) に委託して実施されていた進学適性検査 (SAT) の平均スコアは年々下落傾向にあった。この傾向は

毎年約90万人が受験するアメリカ大学テスト協会の ACT でも同様であった (例えば、ボイヤー, 1984)。特に、1972年以降は大学の志願者構造は安定しており (荒井, 1985), SAT には全高卒者の約3分の1, 大学進学希望者の半分, 年間延べ150万人, ACT には90万人の受験者があった。SAT, ACTの重複受験者も相当数に上るとされる。ちなみに、共通第1次学力試験発足以来、最高となった平成9年度大学入試センター試験の志願者数約60万人と比較すると、2つの試験の実施規模の大きさが実感できる。SAT, ACTともに、主に選抜資料としての高等学校の学業成績を補助する形で利用されているので、厳密には大学の入学試験の一環ではない (天野, 1987) が、高等教育側から見た場合には、上述の現象は志願者の学力レベルの総体としての低下を意味する。

いずれにしろ、このデータはアメリカの教育界に衝撃を与えた。SATは「学校内外の広い生活環境の中で培わ

れた知的能力を測定する」とうたわれていた(荒井, 1985)にもかかわらず, 1983年ころにはいわゆる「ボイヤー報告」, 「危機に立つ国家」などの数多くの教育改革書が現れ, 学校教育の内容に対して大きな改革の気運を引き起こした。その一つが初等・中等教育のカリキュラム改革であり, それに関連した教育評価の強化, すなわち, 州内統一テストの実施である(荒井, 1987)。

アメリカには日本の学習指導要領のような全国的に適用される統一カリキュラムが存在しない(肥田野, 1985)。もちろん, 指導要領に沿った教科書も存在しない。したがって, 初等・中等教育の内容や水準は, 学校区や学校, 個々の教師の力量に負うところが大きい。その結果, 地域差や学校差は我が国以上に大きく, その実態を把握するためにはテストを実施すべきであるという発想が起こってくることは理解に難くない。いずれにしろ, 我が国におけるテストの実施目的があくまでも受験者個人の特性の測定に重きが置かれているのに対し, 教育内容の評価というテストの実施目的はアメリカ合衆国の教育制度を反映したものであると考えられる。

1-2 テストの技術的側面

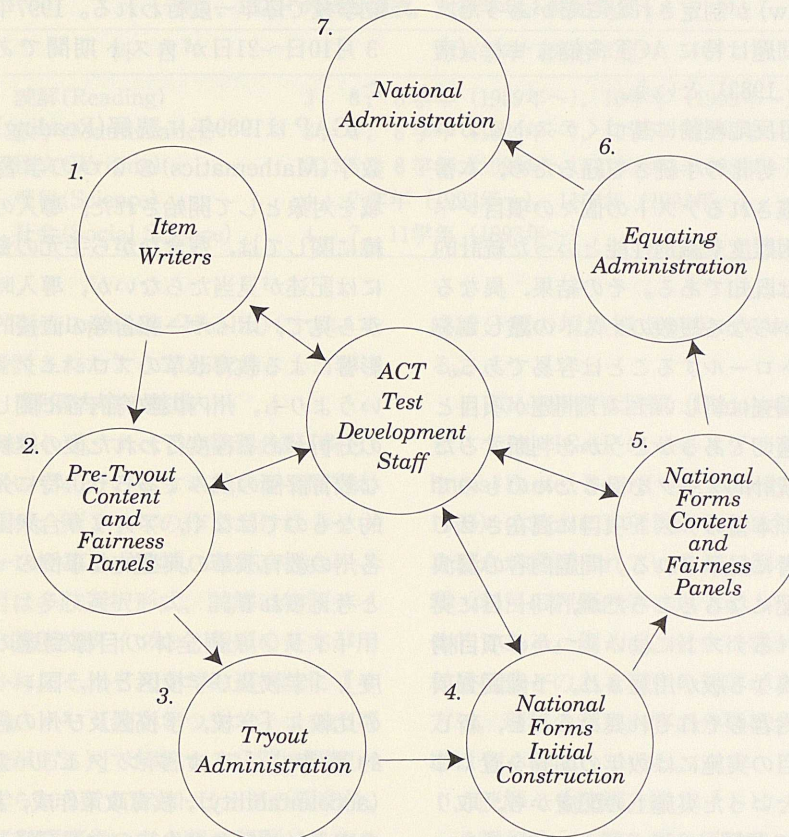
アメリカの1970年代は一種のテストブームの時期であったと同時に, テスト批判も根強く, テスト批判・テスト

論争の時代だと言われた。特に, テスト項目の信頼性, 妥当性のみならず, 文化的バイアスや解釈の問題は論争の中心であった(荒井, 1987)。一方, 理論的立場からのテストバイアスの発見や修正の方法の研究も盛んになった。1980年代には項目反応理論(IRT)の応用が実用段階に達し(池田, 1997), 現在では項目反応理論に基づくテスト設計は客観式テストにおける基礎的, 実用的理論基盤として定着していると思われる。

図1は現在のACTテストの項目作成の流れである。一つの機関の事例ではあるが, アメリカにおけるテスト作成のプロセスと我が国との違いが読み取れる。

大きな相違点の第一は, 「公正さ(Fairness)」についての審査である。ACTの資料によると, テスト開発, 得点の解釈, 公正さへの努力, 受験者への報告という4つの領域に規定があるが, テストの開発者は人種, 性別, 民族的背景, ハンディキャップに対してできるだけ公正なテストを作る努力をすべきである, とされている。多民族国家としてのアメリカの社会的状況が色濃く現れている。

第二は, 予備調査(Tryout)と等化(Equating)の実施である。我が国の入学試験の問題には常に初出のものが要求される。さらに, 我が国の「公正さ」



Dr. Sherri Miller 提供による未公開資料

図1. ACTにおけるテスト作成の流れ

の基準はテスト問題に対する接触機会の平等性に関して厳しい。したがって, 以前実施された問題との重複のチェックは, 作題関係者にとっては頭の痛い話となる。一方, アメリカ合衆国においては, 一般的にはテスト項目はすべてプールされ, 実際の試験にはそこから項目を選んで構成されたテストで行

われる。したがって, 個人に対する結果の通知には多大な神経を使っているにもかかわらず, テスト問題そのものはテストの実施後には回収され, 項目の具体的な内容は実施後も非公開である。テスト批判の流れと「知る権利」の主張から, 1980年にはニューヨーク州でテスト情報公開法(truth-in-test-

ing law) が制定されたことがあった。この問題は特に ACT を悩ませた(南風原, 1983) という。

項目反応理論に基づくテストにおいては、等化の手続きを経るため、本番に実施されるテストの個々の項目レベルで困難度や識別性能といった統計的性質は既知である。その結果、異なる項目からなる複数のテストの難しさをコントロールすることは容易である。予備調査は新しいテスト問題が項目として適切であるかどうかを判断するための統計的データを得るためのものであり、本番のテスト項目に混在させて受験者に呈示される。問題内容の漏洩が心配になるところだが、同一日に実施されるテストにはいくつかの項目構成の異なる版が用意され、予備調査項目の内容もそれぞれ異なること、新しい項目の実施には数年の期間を置くこと、といった実施上の配慮から、取り立てて問題はないということであった。以上のような方略により、項目プールへの新しい項目の補充が一種のルーチンワークとして不断に行われているのである。

2 IGAP

IGAP (Illinois Goal Assessment Program) は、イリノイ州の学校区、学校の教育成果の評価を目的とした悉皆調査であり、州内統一テストである。このテストはイリノイ州のすべての公

立学校で毎年一度行われる。1997年は3月10日～21日がテスト期間であった。

IGAP は1989年に読解(Reading)、数学(Mathematics)の2つの学習領域を対象として開始された。導入の経緯に関しては、残念ながら手元の資料には記述が見当たらないが、導入時期から見て、ボイヤー報告等の直接的な影響による教育改革のプロジェクトというよりも、州内の教育内容に関しての分析がある程度行われた後の継続的な教育評価の試みであろう。特に先進的なものではなく、アメリカ合衆国の各州の教育改革の典型的な事例の一つと考えられる。

「イリノイ州全体の目標到達の程度」、「学校及び学校区と州、国レベルの比較」、「学校、学校区及び州の経年的進歩」、「アカウンタビリティ(accountability)、教育政策作成、学校の向上」、「個々の生徒の学習記録」のための資料とすることが、IGAP 実施の目的であるとされている。

現在の学習分野の実施状況は表1のとおりである。我が国の教科分類と比較的よく対応しているが、中等教育段階においても外国語がないことと我が国の「国語」に当たる分野が読解、作文の2つに区分されていることが特徴と言える。その他、1995年度には美術(Fine Arts)、身体発達/健康(Physi-

表1. IGAP の科目構成、対象学年、開始年度

科目	対象学年 (開始年度)
読解(Reading)	3, 6, 8 学年 (1989年～), 10 学年 (1993年～)
数学(Mathematics)	3, 6, 8 学年 (1989年～), 10 学年 (1993年～)
作文(Writing)	3, 6, 8 学年 (1990年～), 10 学年 (1993年～)
理科(Science)	4, 7 学年 (1992年～), 11 学年 (1993年～)
社会(Social Science)	4, 7, 11 学年 (1993年～)

cal Development/Health) についても検討されたようである。

2-1 評価内容と基準

テストは、すべての学習分野(科目)において40分のセッションが2回行われるが、第3学年の作文だけは1セッションである。数学、理科、社会の3科目は多肢選択形式、読解は物語タイプと説明文タイプの2種類の長文が用意され、それぞれ5つの選択肢に「イエス、ノー(最高3つまでイエスの可能性がある)」で解答する15問の問題が与えられる。また、トピックの親和性、読解方略について2問の問題が与えられる。作文の課題は説得的、説明的、物語的の3タイプから2つ(第3学年は1つ)を選ぶ。

1991～93年にかけてイリノイ公立学校基準体系(Illinois Public School Accreditation System)の一つとしてIGAPの成績基準が作成された。州の教育目標に対して、レベル1は未到達、レベル2は到達、レベル3は卓越であり、それぞれの学習分野、学年ごとに

成績分類の境界得点が定められている。

イリノイ州教育目標 (Illinois Goals) としては、IGAP のガイドブック (Illinois State Board of Education, 1995) の巻末に8項目からなる抽象的な理念が掲げられているが、個々の教科・科目の評価内容は以下のように具体的である。例えば、社会は我が国の科目区分とのある程度の対応がある一方で、理科の測定内容の分類概念の水準は全く異なっている。

- i. 読解: 「読解方略」、「トピックの親和性」、「読み書き経験」。
- ii. 数学: 「数概念と技能」、「百分率、比、割合」、「測定」、「代数的概念と技能」、「幾何的概念と技能」、「データ収集と分析」、「推定/近似」。
- iii. 作文: 「焦点」、「確認/推敲」、「構成」、「用法」、以上の4要素の「統合」。
- iv. 理科: 「理科的な概念と語彙」、「科学技術の発達の意味あい」、「研究

の原理」,「理科的な技法」。

v. 社会:「政治,経済」,「歴史」,「社会科学の基礎知識」,「地理」,「応用」。

2-2 得点スケールと報告の形式

得点スケールについては,読書,数学,理科,社会は0-500点の尺度上に表され,IGAP実施の最初の年の各学年の平均が250,標準偏差が100と標準化されている。これらの学習領域については,テスト項目の等化が行われているので,成績の年度間の推移が尺度得点から直接確認できる。一方,作文については評定者により6-32点の得点が与えられる。作文の課題は毎年異なっているため,年次間の成績の推移を厳密に示すことはできない。したがって,ガイドブックでは,経年変化の得点の解釈については課題の難しさと生徒の成績の両方の要因が反映されていることに注意を促している。

成績の報告様式は学校,学校区用と個人の生徒用の2通りがある。個人用の報告様式は1993年になって初めて作成された。個人の成績は個々の生徒の成績の理解と適切な学習活動の計画のために教師と保護者に通知され,生徒本人には通知されていない。

3 メトリテック

IGAPはイリノイ州教育委員会 (Illinois State Board of Education) と

いう公的機関が実施主体であるが,テストの作成,実施,採点,結果の分析,報告書の作成等の実務を行っているのは公的機関でもなければ,ETSやACTのような大手のテスト会社でもない。

現在,IGAPの実務を請け負っているメトリテック (MetriTech) はイリノイ州シャンペーン市郊外にある常勤職員25名ほどの民間の小さなテスト会社である。社長のサム・クリュッグ (Sam Krug)氏は,私の滞在していたイリノイ大学の出身で,博士号を持っている。人格・能力テスト研究所 (Institute for Personality and Ability Testing)に1969年から1984年まで勤めるかたわら,1982年にメトリテックを設立した。業務内容は,各種テストの開発が中心であり,コンサルティングや調査も行う。ETSやACTは税制面での優遇措置を受けることができる公益法人(not-for-profit company)であるが,メトリテックは通常の企業と同じ立場である。メトリテックのような小規模の民間テスト会社という存在は例外的なものではない。クリュッグ氏によると,同社のライバルとなるような会社はイリノイ州だけでも40~50に上るということであった。

メトリテックでは自社テストを開発,販売する他に,IGAPのような外注のテストプロジェクトをいくつも同時

に請け負っている。例えば,IGAPについては,1992年にイリノイ州教育委員会とテスト開発と技術サポートについて5年契約を結んでいる。ちなみに,契約額は325万ドルだそうである。その他のプロジェクトの契約先は地元だけではなく,シンシナチ,ケンタッキー,ウィスコンシン等の州外にも及んでいる。また,扱うテーマも教育関連のみならず,企業の健康管理や児童保護,ホームレス関連の調査など,多岐に渡っている。

メトリテックはどちらかというと教育関連が得意であるが,広い領域を扱うテスト会社である。一方,医学関連専門などというように,扱う領域がより専門的に特化したテスト会社も存在する。逆に,教育評価テストの大手であるACTでも,近年,職業分析テストに力を注いでいるという話も耳にした。

以上の事例は,アメリカ合衆国においては,問題の作成から実施,分析,報告に至るまでのテストに関わる諸活動が,一つの独立した専門技術であるということが社会的に認知され,日本では見かけないテストの専門業者というものが普通に存在していることを示している。また,行政機関等が実施する公的性格の強いテスト関連事業についても,民間企業が受注して実務を請け負っている。大小多くのテスト関連

会社が経営を成立させるだけの需要が存在するのである。そういう意味では,アメリカ社会は正にテスト社会と言えるであろう。また,テスト会社の扱う内容は教育関連だけに限られたものではない。短い滞在期間の私個人の経験に依った話であるので一般化はできないが,塾や学習教材を提供しているテスト会社というものは聞かなかった。したがって,アメリカのテスト会社という概念は,我が国のいわゆる受験産業と呼ばれるものとは異なる性格のもののようなのである。

アメリカでは我が国には存在しない「テスト産業」とでも呼ぶべき一つの業種が形成されているように感じられた。

文 献

- 天野郁夫 1987 アメリカの入学者選抜 高等教育研究紀要, 7, 73-86.
荒井克弘 1985 SAT (進学適性テスト)の近況について IDE 現代の高等教育, 264, 64-69.
荒井克弘 1987. アメリカにおける教育改革の近況—大学入学問題と教育テスト問題—, 国立教育研究所研究集録, 15, 99-108.
ボイヤー, E. L. 天城勲・中島章夫 (監訳) 1984 アメリカの教育改革—ハイスクール新生の12の鍵 リクルート (E. L. Boyer[1983]. High school:

A report on secondary education in America, Carnegie Foundation.)

南風原朝和 1983 米国における試験と選抜の公平性をめぐって サイコロジー, 41, 44-49.

肥田野直 1985. 高校・大学の教育と入学者選抜 高等教育研究紀要, 5, 7-32.

池田央 1997 心理・教育測定の理論と

技術はいかに発展してきたかーこの20世紀の歴史を振り返るー 立教大学社会学部応用社会学研究, 39, 15-35.

Illinois State Board of Education, Standards and Assessment Division. (1995). Guide to the 1995 Illinois State Assessment.