

2022年11月30日(水)

国立教育政策研究所 令和4年度教育研究公開シンポジウム

「学力アセスメントの動向と展望～CBT化に向けて～」

(オンライン)

米国の公的な学力アセスメントについて

石井英真 (京都大学)

米国における「公的学力アセスメント」とは？

—CBT化を議論する前提として(EdTechのELSI(倫理的・法的・社会的課題)の一側面を問う)

- SATやACTAP(入試:学習者個人の能力判定)? 全米学力調査(NAEP: National Assessment of Educational Progress)(政策評価:社会集団全体の傾向の調査)? 各州の統一テスト(アカウンタビリティ:学校評価・教員評価・質保証)
c.f. 日本における大学入学共通テスト、教育課程実施状況調査、全国学力・学習状況調査の調査目的の違い
- 1980年代以降、米国で「公的学力アセスメント」「学力テスト」として議論の焦点となってきたのは各州による統一テスト(悉皆テスト)とそれを軸にした学力向上(アカウンタビリティによる質保証)政策
- 大規模学力テストの設計においては、その目的やそれが埋め込まれている政策的文脈をふまえた議論が不可欠である。

全米的な学力向上政策の展開

1983年 『危機に立つ国家』(連邦教育省長官の諮問委員会による報告書)の発表

- =アメリカの子どもたちの学力が危機的状況にあることを報告。
- 州知事主導による改革が始まり、州の教育への関与が強まる。
学力向上が国家的な関心事となる。

1989年 「教育サミット」の開催

- =ブッシュ大統領の呼びかけで全米の州知事が一堂に会して、「教育の国家目標(National Education Goals)」(ハイスクール卒業率の向上、数学・理科における世界最高水準の学力の達成などの6項目)を設定。

1991年 『2000年のアメリカ:教育戦略』

- =2000年までに「教育の国家目標」を達成すべく、スタンダードの設定などの実行策が提案された。
- 連邦政府のイニシアチブによる教育改革のはじまり。
州レベルでのスタンダード運動が全米規模で展開される。

2002年「一人の子どもも落ちこぼさない (No Child Left Behind: NCLB)」法の制定

1965年初等中等教育法の改正法であるNCLB法は、不利な条件に置かれている子どもとそうでない子どもとのギャップを12年間という期限付きで埋めることを目的として、補助金交付の条件として下記を各州に求めた。

- すべての子どもに適用されるスタンダードの設定、および、それに基づく州統一の学力テストの開発・実施（第3学年から第8学年までの各学年、第10学年から第12学年については、いずれかの学年で最低1回）。
- 各学区、学校に目標への到達度に関してアカウンタビリティを求め、その結果に応じて報奨や制裁を与える。複数年にわたって年間目標を達成できないと、その学校には段階的な措置が講じられ、最終的に教職員の入れ替えもなされる。

2010年 「共通スタンダード (COMMON CORE STATE STANDARDS)」の発表 (全米州知事協会 (National Governors' Association: NGA) と州教育長協議会 (Council of Chief State School Officers: CCSSO) を中心に開発された)

言語技術と数学に関して、高校を卒業するとき生徒が学習しておくことが期待される内容や能力、および、小学校から高校の各学年において学習しておくことが期待される内容を示している。全米のほとんどの州で採択されている。

スタンダードに基づく教育改革 (standards-based reform) の世界的な展開

• 教育の規制緩和や地方分権を進め、各地域や学校の創意工夫を促す一方で、学校教育に求める目標内容と達成水準(何をどの程度できるようになればよいか:スタンダード)を明確化し、学力テストなどによってその成果を検証したり、アカウンタビリティ(説明責任)を果たさせたりしながら、学校教育の質を保証しようとする(遠隔統治構造)。

↑ テスト準備教育をもたらすものであり、教育を単純作業化・規格化し、教育における競争主義と格差拡大を助長するものといった批判がなされてきた(佐貫・世取山, 2008)。

• 学力テストの活用の仕方については、さまざまな形がある。アメリカのように、テスト結果による制裁が規定されていて、結果に基づいて教師や学校が厳しく値踏みされる国もあれば、オーストラリアのように、結果は公表しても支援的な介入を重視している国もあれば、イタリアのように、悉皆調査を行いつつ全国的な学力水準を確認するためにサンプルを抽出・分析し、あとは学校の自己評価に役立てさせる国もあれば、スウェーデンのように、学校間の評価の質の調整や学力観の共有に役立て、教師の自主的な活用を尊重している国もある(田中, 2016)。

標準テスト中心で競争主義的なアカウンタビリティ政策へのオルタナティブの模索

○標準テスト批判とパフォーマンス評価の誕生

=多肢選択などの客観テストは、現実世界と切り離された無味乾燥な文脈で、断片的な知識・技能を問うものである。交通法規や運転上の留意点をいくら知っていても、上手に運転できるかは実際に運転させてみないと分からないのと同じように、客観テストでは本物の学力は測れない。豊かに考える授業をしながら評価では知識・技能(見えやすい学力)しか問われない、というミスマッチを解消し、目標・指導・評価の一貫性を確立する必要性。

- 教室の外側のハイステイクスな標準テストによる、トップダウンで中央集権的で、学力競争を志向する、説明責任による目標・評価システムに対して、教室の内側の教師の判断を信頼する真正の評価と形成的評価による、ボトムアップで地方分権的で、学力・学習の保障を志向する、応答責任による目標・評価システムというオルタナティブも模索されてきた。

例) 米国における「真正の評価 (authentic assessment)」をキーワードにした教育的で民主的なアカウンタビリティ・システムの模索、「学習のための評価 (assessment for learning)」をスローガンとする、米国のネブラスカ州、スコットランドやウェールズ、カナダのオンタリオ州などでの教室での教師の評価を尊重する評価システムの構築(石井, 2021aなどを参照)。

表. 学校改革の二つのアイデア (出典: Gallagher 2007, 29、石井,2020,)

アカウンタビリティ	関与
<p>企業モデル 学校改革 改革の障害物としての教師 一方通行の関係 生徒の学業達成 テストに基づく 標準化 厳格な父親の道徳 トランザクション トップダウン 統制の行使 ハイステイクス 分不相応な不信 競争 服従 学習の評価(のみ) 単純さを要求する</p>	<p>民主主義モデル 学校改善 学校改善のリーダーとしての教師 相互関係 生徒の学習 豊かな情報に基づく評価 スタンダード 共有された責任性の倫理 インタラクション ボトムアップ(内から外へ) 能力の構築 ハイインパクト 受けるに値する信頼 協働 関与 学習のための評価(も) 複雑さを受け入れる</p>

全米学力調査（NAEP）（1969年開始）の変質

- 第一期（60～70年代） 政策立案に資するデータ供出ツールとして

Main NAEP（主調査：社会的要請の高い教育課題への取り組みを課題とし、最新の教育測定技術の適用と開発を行う）とLong-term Trend NAEP（動向調査：時系列的な学力変化を調べる）を定期的実施。抽出調査で、生徒個人や学校単位のデータにはアクセスできず、州別でなく地域別に限定してデータを収集。

- 第二期（80～90年代） 全米学力調査自身に対する説明責任

連邦政府の教育予算に対する説明責任を果たす「より敏感な」ものにすべく、項目反応理論等の導入による調査デザインの洗練、および州別調査の解禁がなされた（スタンダードに基づくアカウントビリティ政策との連結の始まり）。

- 第三期（2003年以降） 州の説明責任を果たすツールとして

NCLB法で州や学区が毎年作成する評価報告書において、州ごとに比較可能なデータとしてNAEPの州別調査のデータの使用が記されたことで、全州調査となった。

（木村，2008の整理による）

日本の全国学力・学習状況調査(全国学調)の性格

- 日本の全国学調は、改革や取り組みの成果(事実)を検証する役割や、出口(アウトプット)管理の方法としてよりも、実質的には、学習指導要領の学力観の趣旨を伝達し改革を先導・牽引する理念提示機能や、入口(インプット)管理を強化する機能を果たしている(石井, 2021b)。
- 全国学調実施前から、地方自治体レベルでも多くの悉皆の学力調査がなされてきた。しかし、一部の自治体を除けば、保護者や地域住民などへの、市町村別、学校別の学力調査の結果の公表は限定的である。結果の公表が認められる中、競争主義や成果主義を強化する方向で調査結果の利用が広まることには注意が必要である。ただ、英米と比べれば、現時点で日本の全国学調は必ずしもハイ・ステイクスなものではない。
- 日本において全国学調は、英米ほどには、教育の結果を厳格に管理・統制するものとはなっていない一方で、その結果によって上意下達の傾向や授業のスタンダード化が進行しないよう注意が必要である。

CBT化に向けた論点提示

- テストの目的によってCBT化や項目反応理論の導入の可能性や望ましさは変わる。特に、項目反応理論の導入に際して、調査問題が原則的に非公開になる点、問題バンクの作成やデータ管理のコスト、および、問題漏えいのリスクなどが大きな課題となる。過去問で勉強するなどの日本のテスト文化との関係、あるいは、学習指導要領の改訂のたびに問題バンクを変えることになるのかといった日本の教育制度との整合性も検討が必要だろう。
- 調査的性格が強くハイ・ステイクスな性格が弱い、経年変化を見る補完調査、あるいは児童生徒や教員や学校への質問紙調査等への利用はしやすいだろう。逆に、入試への導入はハードルが高い(木村, 2022)。また、悉皆の本体調査も、問題公開の難しさやCBT化しやすいものに学力観が限定されることなどをふまえると、理念提示機能、および個々の学校・教師・児童生徒の指導・学習改善機能において工夫が必要だろう。全国学力・学習状況調査の主目的を絞り、個々の学習者の支援のあり方等についても検討したりしながら(MEXCBT等で別途フォロー)、さらに、学力の質と評価方法との対応関係も考えながら、技術的制約によって目的や中身が規定されないよう注意が必要である。

引用・参考文献

- 荒井克弘・倉元直樹編著(2008)『全国学力調査 日米比較研究』金子書房.
- 石井英真(2015)「教育実践の論理から『エビデンスに基づく教育』を問い直す—教育の市場化・標準化の中で」『教育学研究』Vol. 82, No. 2.
- 石井英真(2017)「学校改革とカリキュラム変革の系譜」『岩波講座 教育 第五巻 学びとカリキュラム』岩波書店.
- 石井英真(2020)『再増補版・現代アメリカにおける学力形成論の展開—スタンダードに基づくカリキュラムの設計』東信堂.
- 石井英真(2021a)「カリキュラムと評価の改革の世界的標準化と対抗軸の模索」広瀬裕子編『カリキュラム・学校・統治の理論』世織書房.
- 石井英真(2021b)「学力調査を飼いならすための評価リテラシー」西岡加名恵・石井英真編著『学力テスト改革を読み解く—「確かな学力」を保障するパフォーマンス評価』明治図書.
- 川口俊明(2020)『全国学力テストはなぜ失敗したのか—学力調査を科学する』岩波書店.
- 木村拓也(2008)「2003年以降の全米学力調査の変質」荒井克弘・倉元直樹編著『全国学力調査 日米比較研究』金子書房.
- 木村拓也(2022)「新しい情報技術と大学入試—大学入試のCBT化を巡る『試論』が導き出すパラドクス」石井英真・仁平典宏・濱中淳子・青木栄一・丸山英樹・下司晶編著『教育学年報13 情報技術・AIと教育』世織書房.
- 佐藤仁・北野秋男編著(2021)『世界のテスト・ガバナンス』東信堂.
- 佐貫浩、世取山洋介編(2008)『新自由主義教育改革—その理論・実態と対抗軸』大月書店.
- 田中耕治編(2016)『グローバル化時代の教育評価改革—日本・アジア・欧米を結ぶ』日本標準.
- C. W. Gallagher, 2007 *Reclaiming Assessment: A Better Alternative to the Accountability Agenda*, Portsmouth, NH: Heinemann.