特集

エビデンス・ベースド・ポリシーと教育調査

エビデンスに基づく政策（エビデンス・ベースド・ポリシー）という言葉が，注目を集めている。矢継ぎ早に改革が繰り返されてきたことで，教育現場は疲弊し，情緒的・感覚的な理由ではなく，客観的かつ科学的な根拠に基づいて政策を実行すべきという声が強まっている。この流れは否定しがたいものであり，客観的かつ科学的なエビデンスを収集する方法として，社会調査が重要性を帯びることに異論はなかろう。

しかしそこで取り上げられるエビデンスとは何か，方法論は確立されているのか，エビデンスの内容に偏りはないのか，エビデンス重視の風潮が教育現場にいかなるインパクトをもたらすのか，といった議論は未だ不十分である。本号の特集では，社会調査を教育政策や教育実践のエビデンスとして活用することについて，様々な観点・立場からの議論を紹介する。今後の教育調査のあり方を，見つめ直すきっかけとなれば幸いである。

（中澤 渉・倉石一郎）

特集論文

1

教育政策・教育実践と エビデンス・ベースド・ポリシー

教育現場における調査のあり方を考える

中澤 渉
大阪大学大学院
人間科学研究科 教授

倉石一郎
京都大学大学院
人間・環境学研究科 教授

1 エビデンス・ベースド・ポリシーの流れ

教育現場における調査が注目を集めている。もちろん，教育社会学の分野では，実証データを収集する調査が長年実施されており，蓄積も多い。その調査の手法は〈量的なもの〉と〈質的なもの〉の両方が含まれる。

いわゆる「学校社会学」において，大型計算機の発達で大量データの処理が可能になった1970年代後半から1980年代にかけて，質問紙調査に基づく学校文化・生徒文化研究が流行した。そうした計量分析偏重に対する批判として，「解釈的アプローチ」とよばれる質的方法に基づく調査研究も導入され，大きなストリームを形成してきたと言ってよい（志水，1985；稲垣，1990）。

しかし2000年前後から，日本では「ゆとり教育」と「学力低下」の関連が取り沙汰されるようになる。つまり「ゆとり教育」政策の問題点が様々な角度から指摘されるようになったが，その政策を評価できるまともなデータを，文部科学省は持っていなかった。それだけではない。そもそも政策を実施する前提となる問題の把握自体が，どこまで客観的になされていたのかすら，怪しかった[1)]。このことは強い批判を浴び，後に「全国学力・学習状況調査」を導入する伏線となる。

ここで政策と，統計分析を前提とする質問紙調査が密接に絡み合うこととなる。質問紙調査の方法論は，相対的には固まっており，出てくる結果も数値なので，科学的かつ客観的に映る。学力問題は，人々の関心も高かった。その結果，生徒の生活や意識，学習態度に関する質問紙調査と，学力データを組み合わせた調査分析が多く蓄積された（苅谷・志水，2004；耳塚，2013；志水・髙田，2016；赤林・直井・敷島，2016など）。

この流れと軌を一にして，国際比較学力調査，特に経済協力開発機構（OECD）の実施する「生徒の学習到達度調査（PISA：Programme for International Student Assessment）」が注目を集めた。PISAは，機能主義・人的資本論の立場から教育を捉え，経済発展のために教育水準を高める方法や，その判定を行うスタンダードを確立するためのデータの獲得を目的とする。2000年に始まったこの調査により，OECD各国の平均的な学習到達度が，白日の下に晒されることとなった。ここで高得点をあげたフィンランドの教育は，世界的な注目を集めた（Meyer and Benavot, 2013；Sahlberg, 2015）。

日本に限らず，多くの欧米諸国では，財源不足に悩まされていた。したがって，少ないコストで，最大限のベネフィットを獲得することが

求められた。その嚆矢が，イギリスにおける医療政策である。紙幅の問題から詳細は他に譲るが（国立教育政策研究所，2012；中澤，2018），この過程で「エビデンス・ベースド」が強調され，やがてそれは教育の分野にも波及した。アメリカでは2001年に，「落ちこぼれ防止法（No Child Left Behind Act）」が制定され，学力向上のため連邦政府の権限が強められた。そこでは説明責任が強調され，学力テストの実施と，その結果に基づく教師や学校の評価が厳格に実施されるようになった。

もちろんアメリカやイギリスと，日本の置かれた社会的文脈は異なる。しかし政府の財政に余裕がなく，納税者の視線が厳しくなっているのは共通している。結局，費用対効果を，全く無視した政策を実施することはできない。ここでエビデンスのない教育政策を強く批判したのが教育経済学者であった（中室，2015）。それまであまり教育に関心をもつことの多くなかった日本の経済学者も，教育に注目するようになり，質問紙調査や実験的手法により，実証的データが本格的に収集されることになった。

2 教育現場とエビデンス

(1) エビデンスの捉え方

ここまで読んで，多くの読者は気づかれると思うが，このコンテクストにおけるエビデンス（証拠）とは，主として集合的データ（aggregate data）や，そのデータに基づく分析結果を指している。こうなると，集合的データとそれを集めるような調査が，ケース研究や質的研究より科学的だとか，客観的だとか，優れているという誤解を招く恐れがある。しかし，いかなる調査法を採用するかは，もちろん問いに依存するのであって，一般論として量的調査と質的調査の優劣を論じることに意味はない。

エビデンス・ベースド・ポリシーの文脈で言及されるエビデンスが量的データに偏っているということは，そのエビデンスを確率論的（stochastic）に解釈しなければならないことを意味する。集合には，一定のばらつき（分散）が存在し，分析結果は白か黒かというような解答ではなく，集合全体からみた蓋然性が示されるに過ぎない。これは，社会科学の計量分析に対する一般的な注意点を述べたスタンリー・リーバーソンによる古典でも述べられている通り，統計分析の結果は集合を構成するそれぞれの個人の行動を予測するものではなく，集合全体の傾向として解釈すべきものなのだ（Lieberson, 1985 : 94-99）。

この点は，エビデンスを教育実践に生かそうとするとき，注意が必要である。私たちは，普段の生活では，自分を独立した一個人と見なすし，他人についても同様に一人の個人と見なしている。そこである教育実践を実行してみて，実際に効果があったかなかったか，という解答は，個人によりバラバラである。

個人からみて，エビデンスに基づいて効果が示された方法を採用して，テストの点が上がれば，効果があったと喜ぶだろうし，上がらなければ効果がなかったと腹を立て「そのエビデンスは間違っている」というかもしれない。当人の自己認識では「当たった」か「はずれた」かの解答は，二つに一つであり，○％当たったとか，×％外れたという状態はあり得ない。しかし確率論的に見れば，エビデンスに対して不服に思う人が一定程度生じるのは必然である。集合全体の蓋然性を示す数値を，個々の人物に当てはめてしまうから，認識のズレが生じるのである。

政策立案に絡めて，量的調査が多く用いられるのには，それなりの理由がある。政策は社会にかかわることで，全体像を把握しておく必要がある。そしてその分布や，変数間の関連から，どこに問題が集中しているかを発見しなければならない。限られた資源を投下するならば，ど

こに投下するのが最も社会に資するのかを判断しなければならない。

ただ，そうした全体像を把握するだけでは不十分なこともある。質問紙調査でも，最近はパネル調査が流行し，時間的変化を追うことが可能になってきてはいるが，パネル調査のコストは非常に大きい上，入れられる変数に限界があり，観察法ほど詳細な変化を捕捉することはできない。また政策的にターゲットとなりうる社会的マイノリティの実態把握も，マイノリティである（少数者である）がゆえに実態が知られにくい。その上，質問紙調査では数を集めにくく，統計分析において大きな誤差が生じ推定が困難になる。そもそも，質的な情報なしに，有効な質問紙を作成したり，統計分析の結果を解釈するのは不可能であろう。

教育社会学の，特に教師の社会学では，教員の多忙な勤務実態が長い間議論されてきた（油布，2009）[2)]。世間の教師に対する視線も厳しくなり，高度な教育力が求められ，より効果的な教育方法は何か，そうした解答を求める風潮が強まっているようだ。こうした風潮にあって，科学的な装いをまとった「エビデンス」という言葉は，魅力的に映る。しかし確率論的なエビデンスは，個々の子どもに，直ちに適用可能であるとは限らない。

蓋然性が高いという意味では，「なんら根拠のない俗説を当てにするよりはマシだ」とは言える。しかし，効果的だというエビデンスのある教育方法を採用すれば，必ず全員に効果が現れるわけではない。統計分析の結果は，個人の集合の平均的な学力が有意に上昇したことを示しているだけで，個別の子どもの成績の伸びは様々である。教育活動は，人間同士の相互作用であり，その相互作用の結果を予測するのはかなり難しい。生身の人間の行うことであり，機械が実行するのとは異なる。教師と生徒であっても，人間関係上の相性の問題もある。同じ教育者が同じ方法を実行しても，相手によって効果が異なることもあるだろう。

学校教育の現場は，予想もしないようなことが起こりうる。予想もしないようなこと，逸脱的な問題は，そもそも件数が相対的に少ないからこそ「逸脱」と見なされる。統計分析で，そうした逸脱を防ぐ可能性の高い対処法は提案できるかもしれないが，逸脱をゼロにはできない。そしてその逸脱したケースをどうするか，という解答を，直ちに与えるものでもない。そうした個別のケースへの対処は，事例研究の方が様々な洞察を提供するかもしれない。社会調査を専門にする人間には自明のことだが，アンケートや統計分析だけが調査ではない。エビデンス，といったときに，そのエビデンスの範囲をあまりに狭く捉えすぎるのは，むしろ視野狭窄（きょうさく）に陥り，問題の解決を難しくすることもあり得るのだ。

(2) 教育現場での調査が目指すもの

なぜ調査が重要なのか。改めて原点に立ち返って考えてみたい。

社会調査をする人間は，社会の実態をより正確に，客観的に把握したいと考えている。問題解決をする上で，正確な実態把握は必要条件であり，これを満たさなければ正しい解決方法を見つける術はない。したがって社会調査のデータは，できるだけ社会の実情をそのまま反映したものであるべきだ。調査自体がデータを歪めたり，特定の反応を誘発するものであってはならない。その目的を達成するために，社会調査法を学ぶことが求められるのだ。

しかし，これは口で言うほど容易ではない。調査という営み自体が社会的行為であり，そこに生きる人との相互作用を伴う。調査は被調査者の協力があって成立するものであり，強制はできない。このため，一般的な標本調査で，回収率が100%になることはほとんどない。あるとしたら，むしろ不自然ですらある。また，回収率が高い方が，確率的に抽出した標本の分布

を反映しているように感じるかもしれないが，回収率の高さ自体は標本の偏りのなさを保証する根拠にはならない。回収率が高くても，特定の性格や属性をもつ回答者が回答拒否するような傾向があれば，サンプルに何らかのバイアスが生じるからだ。とはいえ，調査拒否の多い（回収率の低い）調査は，その調査の内容や実行方法に何らかの問題がある可能性が高いとも言え，回収率の問題を無視してよいと言いたいわけではない。

ここで強調したいのは，学校調査の特性として，教室配布で質問紙を回収する集合自記式調査の場合，状況的に回答を拒否しにくく，回収率が100%近くになりやすいということである。つまり，任意に声をかけられれば回答拒否したかもしれない人々に，拒否しにくい状況を作って調査に参加させているともいえる。そうした人は，調査に対するコミットメントが低いと予想されるので，仮に回収できても，真摯に回答されていない票が含まれるリスクは高まることになる（近藤，2013）。

また教育の調査では，客観的に知りたい情報を把握するという目的に徹せず，評価や競争という異なる目的に利用したいという誘因が発生しやすい。2018年8月，大阪市長が「全国学力・学習状況調査」の結果を人事評価に反映すると表明したのが，その最たる例である。

全国学力調査の歴史を繙くと，1950年代半ばから，小・中・高校生対象の全国調査が実施されてきた。特に中学校では悉皆が原則とされ，通称「学テ」とよばれていた。しかし当時の報道によれば，学テの結果が勤評（勤務評定）につながるとか，県間の競争が煽られたりした結果，事前のテスト問題の漏洩や，成績下位の生徒を休ませて平均点を上昇させようとしたなどの様々な不正行為が報告されている[3]。

調査はあくまで現状把握が目的であって，そのデータを元に改善策を考える材料とするのだという主旨が徹底されなければ，データを歪める様々な誘因が働くことになる。その誘因は，露骨な不正行為として現れるとは限らない。一般的な社会調査であれば，調査目的を厳しく限定しているはずで，集めたデータを目的外に使用するのは，被調査者との契約に反する重大な倫理違反である。もし調査データを科学的な分析目的以外に，評価や査定などに用いるようなことが行われれば（通常の社会調査では，別の目的に流用することは許されないはずだ），当然データの中身をよく見せようというインセンティブが生じる。教育力を問われる教師や学校は，自らの評価を上げようと，成績上昇の見込める生徒を選別して重点的にトレーニングする「教育トリアージ」や，テストの点数を上げるテクニックの伝授に力を注ぐことになりがちである（Booer-Jennings, 2005；Jennings and Sohn, 2014）[4]。

また，文部科学省が学校評価を推進しているため，各学校が学校活動に関して児童・生徒や保護者などに質問紙調査を実施する例が増えている。ところが，専門的な調査教育を受けていない人が質問紙を作成しているためか，回答形式，ワーディング，質問紙の構造などに問題が多く，仮に学校評価として活用するにしても，それが妥当なのか疑問を抱きかねないものも多い。そのような調査の横行は，むしろ教師の多忙化をますます促進し，学校の混乱を招くだけで，正確な情報ももたらされず，むしろ害の方が大きいだろう。

もしデータの収集がうまくいっても，それを分析可能な形で管理・保存しているのか，十分な分析が行われているのか（基礎集計レベルで終始し，ほとんど分析らしい分析が行われていないケースも多いのではないか）という点は，改めて検証されるべきだ。調査にもコストがかかっているのであり，コストに見合った活用がなされなければ，むしろ調査自体の信用を損ね，調査環境をさらに厳しい状況に追い込みかねない。

3　特集論文について

最後に，本号の特集論文を簡単に紹介しておきたい。

上述したように，教育におけるエビデンス・ベースド・ポリシーの推進には，経済学者が大きく貢献している。そして経済学者の分析も，質問紙に基づく大規模調査データの統計分析という点で，社会学者の行う社会調査や計量分析と共通する部分が多い。そこでまず，現在「日本子どもパネル調査（JCPS）」のプロジェクトを統括している赤林英夫氏に，経済学者の立場から，子どもパネルデータの構築の意義や課題，経済学と他の社会科学との違いなどについて論じてもらう。もちろん，経済学と社会学の主たる関心は異なるが，研究対象や分析手法などに重なる部分も多く，お互い学べることも多いはずだ。

次に，教育におけるエビデンス・ベースド・ポリシーの海外の動向を，社会情緒的スキルに着目して，卯月由佳氏に紹介してもらう。ここでいう社会情緒的スキルとは，中室（2015）の紹介した「非認知能力」にあたるものだ。このスキル（能力）は，個人の性格と峻別できない概念である。個性尊重という教育の理念のもと，教育が性格や人格と峻別できない社会情緒的スキルにどこまで介入すべきか，議論が分かれるだろう。ただそうした議論を行う基礎としても，調査とデータの蓄積は継続されるべきだと卯月氏は述べる。

日本の教育政策で，エビデンスとなる調査として最も注目されたのが「全国学力・学習状況調査」（文部科学省）であることに，疑いはないだろう。その導入には一定の意義は認められるが，一方で悉皆調査であることが前提となっており，膨大なコストがかかっている。川口俊明氏には，「全国学力・学習状況調査」の問題点が以前から指摘され続けているにもかかわらず，それが改善されないまま現在に至る理由，そして今後どうしていくべきかについて，問題提起をしてもらう。

以上3論文は，エビデンスとして，定量的，数値的なデータを念頭に置いている。しかし既に述べたように，このようなエビデンスの見方にはある種の偏りがあると言わざるを得ない。なぜなら，社会科学において，エスノグラフィー，ドキュメント，インタビューなど，数値に還元できない質的データの獲得を目的とした調査も，学問の発展に大きな貢献をしてきたからだ。そこで，主として質的調査を行う研究者からも，エビデンス・ベースド・ポリシーと教育について考察してもらうことにした。

酒井朗氏は，20年ほど前に英国で繰り広げられた「エビデンスに基づく教育」をめぐる論争を手がかりに，その根底には教えるという営みや教師の役割の捉え方をめぐる根本的な差異が横たわっている点を指摘する。特に質的研究の立場からの批判により，表層的なエビデンス論議で見落とされがちな，実践の文脈依存性や教師の自律の必要性といった論点が指摘された点は重要である。そこからは，政策や実践に禁欲的である一方で新たな問題理解の視座を提示するという，教育研究のもう一つの貢献の形が示される。

最後に澤田稔氏は，エビデンス・ベースドのアカウンタビリティ重視の流れを，新自由主義や市場主義に連なるものとして批判するだけではなく，一定程度その流れを受け止めた上で，いかにして公正な学校教育が可能になるのかを探求することが重要だと指摘する。その中で，氏自身が訪問調査を行ってきた米国ボストンの事例を通して，教育の質的側面にも目を配ったエビデンス・ベースド・アカウンタビリティの可能性を考察してもらう。

以上のように，分野や方法論の壁を乗り越えて，様々な立場の論考が収録することができた。本特集が，教育現場における調査と政策や実践との関

係を議論するきっかけとなれば幸いである。

注

1) このあたりの経緯については,苅谷(2003)を参照。

2) 最近では,内田(2017)が,位置づけの曖昧なまま,実質的に長時間労働の原因となっている部活動の問題を告発している。

3) 主なものだけでも,1964年10月29日付の読売新聞が報じた,宗像誠也(当時東大教授)や梅根悟(当時東京教育大教授)らを中心とする「学力テスト問題学術研究団」による調査報告のほか,1964年10月15日,1966年6月26日などの記事で,こうした不正行為が触れられている。なお,日教組を中心に,学テが教育の国家統制を強め,教育内容を決める権利(教育権)を国民から国家の手中に委ねるものになるとする反対運動が起きた。特に,北海道旭川市の中学校で,学テに反対する教師が実力行使によりテストをやめさせようとして,公務執行妨害罪に問われたという「旭川学テ事件」はよく知られている。この事件で,一審は学テが国家による教育の不当な支配であると認め,結果的に文部省も学テの廃止に踏み切った(1966年)。上告審ではこの判決は覆るのだが(1976年),大きな騒動になったこともあり,2000年代に入るまで,全国規模で学力を把握する目立った動きはほとんど見られなかった。

4) もちろん,評価自体を否定するわけではない。しかし評価そのものも,何のために行うのか,評価が全体にとってどのようなポジティブな影響をもたらすのか,それがいかにして可能なのか,ということは問われなければならない。

文献

赤林英夫・直井道生・敷島千鶴編,2016,『学力・心理・家庭環境の経済分析－全国小中学生の追跡調査から見えてきたもの』有斐閣。

Booher-Jennings, Jennifer, 2005, "Below the Bubble: 'Educational Triage' and the Texas Accountability System.", *American Educational Research Journal*, 42(2): 231-268.

稲垣恭子,1990,「教育社会学における解釈的アプローチの新たな可能性－教育的言説と権力の分析に向けて」『教育社会学研究』47: 66-75。

Jennings, Jennifer L. and Heeju Sohn, 2014, "Measure for Measure: How Proficiency-based Accountability Systems Affect Inequality in Academic Achievement." *Sociology of Education* 87(2): 125-41.

苅谷剛彦,2003,『なぜ教育論争は不毛なのか』中央公論新社。

苅谷剛彦・志水宏吉編,2004,『学力の社会学－調査が示す学力の変化と学習の課題』岩波書店。

国立教育政策研究所編,2012,『教育研究とエビデンス－国際的動向と日本の現状と課題』明石書店。

近藤博之,2013,「生徒調査における回答者の非協力的態度について」『大阪大学大学院人間科学研究科紀要』39: 39-56。

Lieberson, Stanley, 1985, *Making It Count: The Improvement of Social Research and Theory*. Berkeley: University of California Press.

Meyer, Heinz-Dieter. and Aaron Benavot eds., 2013, *PISA, Power, and Policy: The Emergence of Global Educational Governance*. Oxford: Symposium Books.

耳塚寛明編,2013,『学力格差に挑む』金子書房。

中室牧子,2015,『「学力」の経済学』ディスカバー・トゥエンティワン。

中澤渉,2018,『日本の公教育－学力・コスト・民主主義』中央公論新社。

Pasi Sahlberg, 2015, *Finnish Lessons: What Can the World Learn from Educational Change in Finland?*, Second Edition. New York: Teachers College Press.

志水宏吉,1985,「『新しい教育社会学』その後－解釈的アプローチの再評価」『教育社会学研究』40: 193-207。

志水宏吉・髙田一宏編,2016,『マインド・ザ・ギャップ－現在日本の学力格差とその克服』大阪大学出版会。

内田良,2017,『ブラック部活動』東洋館出版社。

油布佐和子編,2009,『教師という仕事』日本図書センター。

特集論文

2

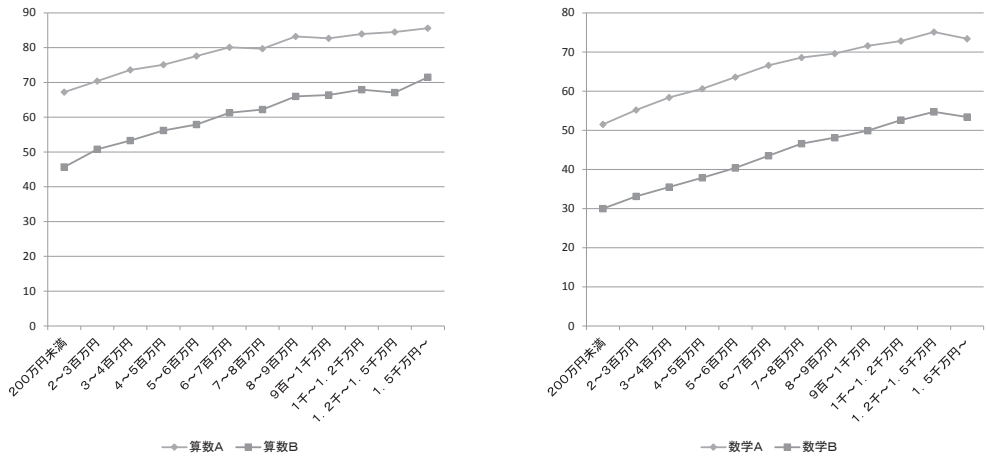
政策評価のために必要なデータとは
「日本子どもパネル調査」の経験から

赤林英夫
慶應義塾大学経済学部 教授

1　はじめに

近年，日本を始めとする先進諸国で，子どもの貧困の増加と，家庭の経済格差から派生した子どもの教育格差が，次世代の経済格差を再生産するのではないかと懸念が広がっている。

図１（浜野，2014）は，2012年の全国学力・学習状況調査を受けた子どもの保護者に対する追加調査に基づき，子どもの平均的学力と世帯所得の関係を示したものだが，中学３年生の最上位の所得層を除き，世帯所得と学力の間に正の相関があることが分かる。この図はメディアで繰り返し取り上げられ，「親の経済的格差の拡大が子どもの教育格差につながり，経済格差が世代間で再生産されている」というエビデンスとして使われている。この図に基づき，経済的余裕のある家庭の子どもだけが，塾や私立に通うなどの質の高い教育を受け，学力が上がり，進学や就職にも有利になると主張され，経済的に不利な環境に置かれている子どもへの学業支援政策の充実に対し，大きな力を持つようになった。近年のOECDの調査でも，子どもの貧困率に関して，日本はOECD加盟国の35カ国中11位



（1）小学６年生　　（2）中学３年生

図１　平成25年度全国学力・学習状況調査（きめ細かい調査）に基づく世帯所得と算数・数学の学力の関係
（縦軸：学力（偏差値）　横軸：世帯税込み年収）
出典：浜野（2014）図表2-1-65「世帯収入（税込み年収）と学力の関係」に基づき筆者作成

で,下位のグループに属することが明らかになっている。

子どもの貧困は,「機会の均等」という誰もが否定できない社会的価値に対する危機につながる。その理由は,家庭の社会経済的条件の違いが,就学以前や,就学後の家庭や学校外での教育機会に差をつけるからだ。そして,教育の剥奪は,成人してからの社会生活,所得や健康の格差と関連することを示す多くの研究がある。

しかしながら,図1が示すのは,すでに生じている子どもの学力格差と家庭の経済格差の間の水平的関係に過ぎない。家庭の経済格差が教育格差にどのていど動的に結びつくのか,そのメカニズムはいかなるものか,どのような政策をいつ実施すれば学力格差の広がりを抑止することが可能なのかについては理論的・実証的にも解明が必要な点が多いのは明らかだ。

特に,貧困は,単に所得が低いだけではなく,その家庭が社会の中で孤立し,情報を与えられず,将来の展望もなく,子育ての気力を失っている状況を同時に生じさせる。その場合には,金銭的所得以外のさまざまな要素がこどもに負の影響を与え,それら要素に対する効果的な政策も異なるはずだ。しかし,政策の貢献を個別に定量化するのは容易なことではない。

そのため世界中で,分野横断的な研究グループにより,こどもの成長過程を,豊富な家庭背景情報や学校情報とともにデータとして残しつつ,それを用いて,家庭背景・教育政策がこどもの成長に与える影響を分析する努力が進められている。一方日本では,経済格差と教育格差の間の相関の認知は進んだが,その発生メカニズムを因果的に明らかにするために必要なデータや手法について,教育政策の場において議論されることは多いとはいえない。

筆者は2010年から,教育行政学者・教育心理学者等と共同で,こどもの追跡調査である「日本子どもパネル調査(JCPS)」を実施してきた。本稿ではその経験を踏まえ,次世代の育成のための政策の評価を目的とした調査と分析のあり方を,経済学の立場から説明する。

次世代の育成のための政策といってもその領域は膨大である。政策を,家庭,特に子どもをもつ貧困家庭に対する支援政策と,学校を通じた教育政策に分けて,本稿では前者に話題を絞る。そして,所得格差が教育格差に繋がる理論的背景を説明した上で,弱い仮定の下で,所得から教育への因果関係を実証的に明らかにするために必要なデータ,変数,そして推計方法について,順に議論する。特に,これらの手順の背景に存在する経済学固有の考え方についても説明を加える。

2 家庭の経済状況が子どもに与える影響 経済モデルとパネルデータの意義

エビデンスに基づく政策策定(EBPM: Evidence-based policy making)には,定量的データが欠かせない。特に,同じこどもを長期間追跡することで得られるパネルデータ・コホートデータが,政策評価の為には有益であることはよく知られている。だが,なぜそれらが有益なのか,理論的な認識が広く共有されているとは言いがたい。経済格差と教育格差の間の因果関係の議論には,経済学的な視点が必要不可欠だ。ここでは,教育の価値と教育行動の関係を説明する中心的概念である〈人的資本理論〉を用いて説明する。

教育を投資として考える

人的資本理論では,家庭・学校を問わず,子どもに多くの教育を与えると,得られた知識・教養や技能は「人的資本」として将来の労働市場での生産性を向上させ,賃金や所得などの便益を増大させると考える。つまり,教育は現時点で費用を支払い,将来の価値増加を期待する投資である。

教育を施すほど将来の子どもの期待所得は上

昇するが，その上昇幅は次第に減少してくるはずだ。たとえば，大卒と高卒の間の生涯所得の差は大きくても，大卒と大学院卒の差はそれほど大きくない。つまり，教育投資を増やしていくと，そのための費用と得られる便益の比[1]は少しずつ減少すると考えられる。これを「教育の限界収益率の逓減」と呼ぶ。

子どもの潜在能力が高いと，教育の限界収益は高くなるはずだ。一方，低所得家庭は，教育の費用を外部から調達するため，利子などの資本コストが発生したり，借入制約に直面したりする。現実の資本市場は完全ではなく，調達金額が増えると，利子率の高いところからも借りる必要があり，所得の低い家庭ほど，追加的に必要な資金調達費用は上昇する。

ベッカー（1976）は以上の概念を用い，家計所得と個人の能力の差により教育格差が生まれ，次世代の経済格差が発生するメカニズムを仮説として提示した。

図2で，横軸は教育投資量を，D1-D3は各々異なる個人の人的資本投資の限界収益率を表す。右下がりの形状は「教育の限界収益率の逓減」を意味し，子どもの潜在能力が高いと，D3からD1にシフトする。S1-S3は経済状況が異なる家計の教育投資の資金調達費用（利子率）を表す。右上がりの形状は，資金調達費用は投資金額が増えるに従い上昇することを意味する。また，所得や貯蓄水準が下がると曲線はS3からS1にシフトする。

親にとっては，子どもへの教育投資の限界収益率が利子率に一致するところまで教育を施すことが合理的となる。たとえば，D2の能力の子どもを持ち，S2の利子率に直面する親は，両者の交点であるE2までの教育投資を行う。それ以上教育を与えても教育による収益が資金調達費用を下回り，投資として価値がマイナスになるからだ。

この図から，社会における家庭の経済状態と子どもの潜在能力の相関の有無が，次世代の経済格差に大きな影響を与えることが分かる。

もし，生まれつき能力の高い子どもが同時に豊かな家庭で育つ「二重の有利」の傾向にあれば，子ども（D）と家庭（S）の組み合わせはD3-S1，D2-S2，D1-S3となり，教育水準はE1，E2，E3の３点に分布する。すなわち，能力の高い子どもほど多くの教育を受ける機会を与えられ，能力の低い子どもほど教育を受ける機会を失う。教育の差が所得の差に直結するのであれば，これは経済格差の拡大・再生産を意味する。

因果識別のためのパネルデータの意義

家庭がこどもの教育達成に与える影響を実証的に明らかにすることは，世代間の社会経済的格差の伝搬メカニズムの解明のためには必須である。しかし，上で議論したように，教育格差を生み出しているのは，親の経済状態と子どもの潜在能力の両方だ。観測できない子どもの能力の差を考慮せず，クロス・セクション・データに基づき親の所得だけで子どもの教育達成度を説明しようとすれば，経済格差が教育格差に与える因果的効果を過大評価する可能性がある。

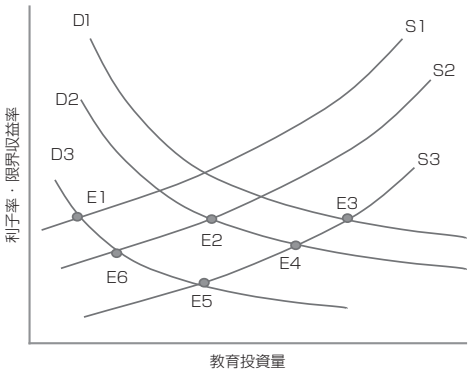


図２　家庭の経済状況と子どもの潜在能力により異なる最適な教育投資量の選択

出所：赤林英夫・直井道生・敷島千鶴2016「経済格差の再生産と教育の役割 —— 子どもと家庭の追跡調査はなぜ必要なのか？」赤林・直井・敷島（2016）所収.

家庭の経済状態と，子どもの学力や成長の間の因果関係を識別するために，現在ある手段の中でもっとも有効な方法が，パネルデータを利用した研究である[2)]。パネルデータを用いれば，一定の弱い条件の下で観測できない属性の影響を取り除き，家庭の経済状況が子どもに与える影響をバイアスなく計測できる[3)]。

図2に戻ると，たとえば，生来の能力がD3で固定された子どもを追跡調査し，親の所得の変化とともに，子どもに対する教育投資量がE1, E6, E5と増加したことが観測できたとする。これは，所得の変化が資金調達費用をS1からS3にシフトさせたことで生じているため，親の経済状態が教育投資量に与えた因果的効果だと解釈できる。つまり，世帯の収入や子どもの学習のための支出，そして子どもの学力や非認知能力等のアウトカムなどの情報を，一定間隔で正確に収集したパネルデータがあれば，子どもの潜在能力の影響を除去し，所得が教育投資に与える純粋に因果的影響の推計を行うことができる。これが計量経済学でいう「固定効果モデル」である。

このように，子どもの学力や心理の変化の背景にはどのようなメカニズムが存在しているのか，因果関係に踏み込んだ分析を行うためには，同じ子どもを追跡したデータが必要だ。そして，追跡データがなければ，ある学力や心理状態の子どもが，どのような青年期を送り，その後どのような成人になるのか，やはり知ることは不可能である。

3 「日本子どもパネル調査」と家計所得情報

「日本子どもパネル調査」開始の経緯と特徴

筆者を中心とした研究グループは，2010年より，慶應義塾大学パネルデータ設計・解析センターにおいて，子どもの学力と心理，さらに家庭環境の詳細な情報を親と子から収集し，小学1年生から中学３年生まで，継続的に追跡調査を実施する我が国で初めての全国調査「日本子どもパネル調査（JCPS: Japan Child Panel Survey）」を設計・実施し，得られたデータを分析してきた（赤林・直井・敷島，2016）。

JCPSは，日本の成人代表サンプル（サンプリング時点）を毎年追跡調査する，日本家計パネル調査（JHPS: Japan Household Panel Survey），および慶應義塾家計パネル調査（KHPS: Keio Household Panel Survey）の付帯調査として出発した。JCPSの設計に際し，すでに存在する豊富な世帯・家計情報を有益に活用するために，我々は，米国において，子どもの教育成果（アウトカム）と家庭背景との関係に関する研究に飛躍的に貢献したNational Longitudinal Survey of Youth（NLSY）――Child Supplement（CNLSY）を参考にした。NLSYは1979年以降，原則毎年，対象の女性に所得や生活に関する多彩な調査を行っていたが，1986年以降，2年に一度，対象者の子ども全員に，発達や行動・学力等の調査を実施している。

CNLSYは，発達心理学，教育心理学の研究者により設計されたが，結果的に，教育経済学の研究にも大きな貢献をした。たとえば，Carneiro and Heckman（2004）は，この調査を用いて，所得階層間での学力格差は幼児期から拡大していることを発見している。同時に，2000年代後半から，家庭背景とアウトカムの関係の国際比較に対する関心が急速に高まっていた（Smeeding *et al.*, 2012）。そこで我々は，JCPSの主たる目的を，「家庭背景と子どものアウトカムの関係の国際比較研究への貢献」と設定し，CNLSYを可能な限り参考にし，日米間の比較可能性をあるていど担保しようと考えた。

一方，JCPSは，すでに実施されていたJHPS/KHPSの調査方法や研究予算の制約を踏まえる必要があった。例えば，JHPS/KHPSの調査開始時点では，調査協力者には，子どもも調査対象にすることの同意を取っていなかった。そこで，JHPS/KHPSの回収率への影響を最小限に抑え

るために，JCPSへの協力はあくまで任意としたため，調査の都度同意をとる必要が発生した。また，調査はJHPS/KHPSと同様に，郵送により2月に行った。

JCPSは2部の調査票から構成され，第１部は子ども票，第2部は親票で，JHPS/KHPSの対象者は，子どもひとりにつき親票と子ども票を一部ずつ回答する。子ども票では，主に，国語と算数／数学の基礎学力，演繹的論理形式に基づく推論能力，生活の充足感（QOL），放課後の過ごし方や学びについての項目が。一方親票では，通学状況や教育環境，学校種別（公立・私立），学級規模，通学時間，受験，勉強時間，学習環境，学校外教育，教育支出，就学前教育，子育て方法，子どもの社会性などの項目を収集した。

教育のパネルデータとしてのJCPSの特徴は，(i) 全国をカバーする調査対象，(ii) 多様なアウトカムの測定，(iii) 家計調査を通じた豊富な家庭背景情報，(iv) 研究者であれば誰でもアクセス可能な研究インフラ，といえる。

(i) は，日本全体を対象とした現状分析・政策研究のためには必要な条件である。

(ii) については，国際比較を意識し，社会性やQOLなどの国際的に標準化された測度を導入し，子どもの「非認知能力」が果たす役割を調査している。算数／数学・国語に関しては，学年ごとに異なる問題を用意しているが，昨年度実施の研究で，項目反応理論に基づく垂直等化を実施し，学年間の学力を共通の尺度で比較可能にしつつある。

(iii) は，第1に，自治体を通じた公立学校をベースにした調査と比較すると，私立進学や転居・転校によるサンプル対象の脱落を抑止できる点が最大の利点である。第2に，学校を通じた調査では困難な兄弟姉妹の特定や，親の就労状態や所得源，介護の必要な家族の存在など，詳細な情報の取得を可能にしている。たとえば，親の健康状態や介護の発生など，子どもの教育とは直接関係のない変数は，世帯が利用可能な経済資源や時間資源の変化を通じて，子どもの教育に影響を与えることから，家計と教育の因果分析のために有益である。一見「教育と関係ない」変数がデータセットにどれだけ用意されているかも，政策の因果分析実施の際のカギになる。CNLSYも豊富な家庭背景情報を持つことで，家庭が子どもに与える因果的影響の分析を可能としてきた[4]。

最後に，(iv) は，研究の自由という観点から重要である。例えば，全国学力調査を始め，国や自治体の調査報告書で，性別ごとの分析が示されることはほとんどない。性差を明らかにすることで「偏見を助長する」，「政策で変えられない部分に注目してどうするのか」という批判を恐れるのかも知れないが，理科系を目指す女性が少ない背景の分析など，研究課題は多いはずである[5]。

経済支援政策効果の因果分析に必要な情報

経済学は所得の計測にこだわりがある点で，「社会経済的背景（SES）」で議論することの多い他の社会科学と異なる。それは，政策がアウトカムに与える因果的効果の背景に，家計が直面する予算制約の外生的変化を想定するからである。

予算制約を決定づけるもっとも重要な要素は所得である。実際，政策が家庭に対してできることは限られている中で，最も簡単かつ伝統的に行われてきた方法は，所得を補填することであった。そのため，経済学では，そもそも所得という経済的資源はどれだけ子どものアウトカムや教育支出に因果的に影響するのかについて多くの研究の蓄積がされてきた。

皮肉なことに，その過程で明らかになってきたのは，予算制約の識別は一筋縄にはいかない，ということだ。

その理由の第１は，所得には内生性の問題があるからだ。多くの世帯にとって，主要な所得源は勤労収入であり，それは働き方の選択によっ

て決定される「内生変数」である。働き方は，所得だけでなく，家庭内での時間配分にも影響を与える。つまり，所得と家庭内時間は同時に変化して子どもに影響を与える可能性がある。したがって，勤労所得よりも外生的であると思われる不労所得（資産所得，移転所得）をデータで区別できることが理想である。

第2に，予算制約を決定するのは，税引き前の所得ではなく，可処分所得である。可処分所得の計算のためには，課税所得を確定する必要があるが，そのためには，本人の扶養家族などの世帯構成や所得源ごとの収入金額が正確に把握できなければならない。

第３に，所得の詳細を尋ねるほど，回答者に負担がかかる。そもそも所得を正確に記憶している人は少ないため，計測誤差による推計のバイアスは無視できない。例えば，子ども手当の効果を実証したいと思っても，その金額を正確に記憶し，世帯の所得金額に含めて回答できる人は多くはないであろう。そうである場合，子どもの数や年齢などの外形的情報を用いて，その世帯が受け取る子ども手当の金額を推計できるようにした方がよい。そのためには，課税所得の推計値に加え，世帯の詳細な情報が必要となる。

JCPSの親世代を対象としたJHPS/KHPSでは，回答者に，事実上税務申告書と同様の記載を依頼しているのは上記の理由からである。それは，政策効果の検証にはふさわしい設計とは言えるが，回答者には相当な負担となっているはずだ。経済学者以外の社会科学者（たとえば社会学や教育学）は，所得や収入に代替する変数を組み合わせて，家庭の文化水準やSES変数を構築することも多い。あるいは，回答者の負担の軽減と回収率向上のために，世帯所得をカテゴリーに分けて尋ねることが多いが，それでは計測誤差の問題が大きくなり，所得補填政策の効果の検証に使うことは困難である。ここは，JHPS/KHPSとJCPSが直面するジレンマであるのは間違いない。

4 所得が子どもに与える因果的効果の実証研究

海外での研究例

家庭の経済格差と子どもの教育格差の間の因果関係を解明するために，海外の多くの研究では，家庭背景，教育支出，そして子どものアウトカムを計測したパネルデータが利用されている。初期の研究では，単純な固定効果モデルにより，子どもや家庭の観測できない要因の影響を除去した上で，所得の変動が学力などのアウトカムに与える影響を分析している（Duncan *et al.*, 1998; Blau, 1999）。近年の研究では，所得の内生性の問題を回避するために，外的要因による所得の変動や，所得補填政策の対象者が機械的ルールで異なることを操作変数（すなわち，教育に直接影響を与えない外生的変動）として推計を行うようになってきた。 別の言い方をすると，勤労収入ではなく不労所得の偶然の変動を利用して，所得が子どもに与える影響を推計している。

たとえば，Løken（2010）は1970-80年代のノルウェーでの海底油田発見がもたらした所得の増加を操作変数としたが，所得上昇が子どものアウトカムへ与えた効果は発見できなかった。一方，Akee *et al.*（2010）は，ノースカロライナ州のネイティブアメリカン保留地におけるカジノ開設がネイティブアメリカン世帯に直接の所得上昇の恩恵をもたらしたことを操作変数として利用し，所得上昇は子どもの高校卒業率向上や犯罪への関わりを減らすなどの効果があったとしている。税制や生活保護制度の変更も外生的な所得の変動をもたらすことから，この問題の解決に利用される。Dahl and Lochner（2012）は，米国における給付付き勤労所得税額控除（EITC）の変更に従った所得上昇の恩恵を受けた世帯の子どもの学力が向上したとしている。Gregg *et al.*（2006）は英国における生活保護制度の変更が低所得世帯における子どものための支出を増加させ，所得階層間の教育格差を縮小させる効

果があったとしている。

JCPSを用いた筆者らの研究

最後に，このテーマに対して，我々が最近発表した研究を紹介する（Naoi *et al.*，2017）。

我々は，JCPS2010-2014を用い，近年に頻繁に変更された子ども手当・児童手当を可処分所得に対する外生的ショックと考え，所得が子どもに対する教育支出や項目反応理論で垂直等化された学力尺度に与えた因果的効果を推計した。特に，手当の金額は，子どもの年齢や出生順で外形的に決定されるため，それらの支給条件と金額の変更は，予想不可能な形で外生的に変化する。本研究では，2009年から2013年にかけて生じた児童手当（子ども手当を含む）制度の変更により，予想外の形で変更された金額部分を操作変数とし，それが可処分所得の変化を通じて，子どもの学力や教育支出に与えた影響を推計した。

過去の海外の研究では，子どもごとの教育支出をアウトカムとして分析している例はなく，また，子どもに対する教育支出と子どもの学力への影響を同時に推計している例もほとんどない。従って，これらをすべて同じデータセットで推計できることが，国際的に見てもJCPSの利点だと考えている。

分析の結果，最小二乗法（OLS）や階差モデル（FD）の推計によると，家計所得と子どもの学力，教育費支出は正の相関を持つことが確認された。しかし，階差操作変数法（FDIV）によると，家計所得は学力に統計的に有意な影響を与えていないことから，OLSやFDで観察された結果は因果的効果を示していないことが示唆された。ただし，FDIVの下でも，家計所得は教育費支出に正の影響を与えていることが示された。

この研究では，世帯の所得情報や世帯構成がパネルデータで詳細に把握できていることが大いに生かされているが，データの制約に由来する限界もある。例えば，今回用いたのは5年分のデータにすぎず，所得が学力へ与えた効果が発見できなかったといっても，それはあくまで短期的な影響であり，長期的な効果については分析対象外である。

5　結語

客観的に判断する材料に乏しい中では，学習支援や保護者への経済支援を実施するにしても，どれほどの規模と期間で実施すれば良いのか，何を優先すればよいのか，などの明確な指針を作ることは難しい。実証的指針がない中での政策立案は，必然的に「あれもこれも」と総花的となり，限られた公的資金の食い合いとなるため，政策現場や教育現場に負担がかかる割には，効果があがりにくい。EBPMは，教育政策・家庭支援政策でも必須であるのは間違いない。しかしながら，政策現場や統計担当者の間では，EBPMのためにはどのようなデータが必要か，どのような変数が必要か，データはどのような状態で保管されていなければならないかという理解が進んでいるとは言えない。いざ「EBPMをやろう」となったところで，利用可能なデータ，意味のある項目や変数が一つもない，という現状をしばしば耳にする。

そのような状況を早く断ち切るために，JCPSを始めとして，分野の垣根を超えた研究者が国際的な視点で収集・構築している世界各国のデータを，多くの人の参考にしてもらいたいと感じている。

本稿では，家庭と子どものアウトカムの関係に話を絞った。それでは，学校で収集されるデータに，政策のために必要な情報は含まれているであろうか。例えば，教員の労働時間と子どものアウトカムに因果関係はあるか，それを検証するためのデータや変数はどうあるべきか，きちんと議論され，保管されているであろうか。

JCPSは限られた予算とさまざまな制約の中で実施されているため，欠点も少なからず存在

する。例えば，別調査の付帯調査であるため，サンプル数を増やすことが困難であること。また，回収率が高いとは言えず，学校に関することなど，センシティブな情報を収集することが困難であることだ。その意味で，JCPSは，最近少しずつ広がりつつある自治体が主体となって実施する，学力・学習状況の調査と補完的な関係にあるだろう。

今後の子どもの調査の主流はどうなるか。世界的に見ると答えは明らかで，それは学校と世帯の「行政データ：administrative data」の接合と活用である。この点でも日本の動きは世界から大きく取り残されている。EBPMは，行政にも，政策評価のための強い意志，必要なデータや変数の理解，そして利用可能な形でのデータ保管ルールの設定と運用がなければ進まない。そのためにも，研究者は積極的に行政との連携や情報提供を進めるべきであろう。

注

1）費用と便益の比から１を引いたものを「教育の収益率」と呼ぶ。

2）理想的なランダム化を通じた社会実験を実施することができれば，政策効果の因果推論のためにパネルデータは必ずしも必要ではなく，事後的なアウトカムデータを１回だけ収集できれば良い。また，回帰切断分析などの手法を用いることで，クロスセクションデータのみで因果分析は可能である。しかしながら，これらも分析に際し，ランダム化や切断が理想的に実現しているかの確認のために，何らかの事前データは必要である。

3）詳細は，赤林・直井・敷島（2016）の巻末付録を参照のこと。

4）この点は実は非常に重要なのだが，政府の調査設計担当者に理解してもらうことは非常に難しい。調査の「政策目的」が特定されればされるほど，一見「目的と関係ない」変数を含めることは不可能になる。

5）赤林・野崎・湯川(2018)はJCPSを利用して，学力達成や学校外教育の支出の男女差を家庭の経済状態や親の学歴別に分析を行っている。

文献

赤林英夫・直井道生・敷島千鶴（編著），2016，『学力・心理・家庭環境の経済分析－全国小中学生の追跡調査から見えてきたもの』有斐閣。

赤林英夫・野崎華世・湯川志保，2018，「日本における教育投資と教育成果の男女間格差：日本子どもパネル調査を用いた分析」Keio-IES Discussion Paper Series. DP2018-004.（https://ies.keio.ac.jp/publications/9195/）。

内閣府，2014，『子供の貧困対策に関する大綱』。

浜野隆，2014，「家庭環境と子どもの学力」『平成２５年度「学力調査を活用した専門的な課題分析に関する調査研究」平成２５年度 全国学力・学習状況調査（きめ細かい調査）の結果を活用した学力に影響を与える要因分析に関する調査研究』所収，国立大学法人お茶の水女子大学（http://www.nier.go.jp/13chousakekkahoukoku/kannren_chousa/hogosya_chousa.html）。

Becker, Gary, 1967, *Human Capital*, 2nd ed.（佐野陽子訳，1976『人的資本 ― 教育を中心とした理論的・経験的分析』東洋経済新報社）.

Akee, R. K. Q., W. E. Copeland, G. Keeler, A. Angold and E. J. Costello, 2010, "Parents' Incomes and Children's Outcomes: A Quasi-Experiment Using Transfer Payments from Casino Profits", *American Economic Journal: Applied Economics*, 2(1): 86–115.

Blau, David M., 1999, "The Effect of Income on Child Development", *Review of Economics and Statistics*, 81(2): 261–276.

Carneiro, Pedro and James Heckman, 2004, "Human Capital Policy", J. Heckman and A. Krueger(eds.), *Inequality in America: What Role for Human Capital Policy?*, MIT Press.

Dahl, G. B. and L. Lochner, 2012, "The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit",

American Economic Review, 102(5): 1927–1956.

Duncan, G. J., W. J. Yeung, J. Brooks-Gunn and J. R. Smith, 1998, "How Much Does Childhood Poverty Affect the Life Chances of Children?" *American Sociological Review*, 63(3): 406–423.

Gregg, Paul, Jane Waldfogel and Elizabeth Washbrook, 2006, "Family Expenditures Post-Welfare Reform in the UK: Are Low-Income Families Starting to Catch Up?" *Labour Economics*, 13(6): 721–746.

Løken, K. V. , 2010, "Family Income and Children's Education: Using the Norwegian Oil Boom as a Natural Experiment", *Labour Economics*, 17(1): 118–129.

Naoi, M., H. Akabayashi, R. Nakamura, K. Nozaki, S. Sano, W. Senoh, C. Shikishima. 2017, "Causal Effects of Family Income on Child Outcomes and Educational Spending: Evidence from a Child Allowance Policy Reform in Japan". Keio-IES Discussion Paper Series. DP2017-026 (https://ies.keio.ac.jp/publications/8579/).

Smeeding, T., R. Erikson and M. Jantti, 2012, *Persistence, Privilege, and Parenting: The Comparative Study of Intergenerational Mobility*, Russell Sage Foundation.

特集論文

3

エビデンスの広がりと問われる教育政策
社会情緒的スキルの教育と調査をめぐる欧米の動向から

卯月由佳
国立教育政策研究所国際研究・協力部 主任研究官

1 はじめに

日本でエビデンスに基づく教育政策の立案が目指されるなか，「何が有効か（What works?）」に関するエビデンスの蓄積が急務である（日本では特に遅れている）という認識が強まっている。一方，そうしたエビデンスを教育政策のどのような目的に対してどのように用いるか，そもそも教育政策にとってどのようなエビデンスが必要かを問う議論は，蓄積が遅れているばかりか，その必要性についての認識も不十分なままである。しかし，近年の欧米の動向を見ると，「何が有効か」のエビデンスの蓄積だけでなく，エビデンスの利用の前提となる規範や価値に関する議論がますます重要になっていると考えられる。本稿は，この課題について，欧米における社会情緒的スキル（social and emotional skills）の教育と調査の動向を取り上げ，考察する。

社会情緒的スキルは，従来注目されてきた認知的スキル以外のものとして非認知的スキルと呼ばれることや，社会的スキル，心理社会的スキル，ソフト・スキルなどと呼ばれることもある。より具体的には後ほどレビューするが，およそ粘り強さ，自制心，好奇心，楽観的思考，自信，協調性などを指すとされる。このスキルが，学力や認知的スキルとともに，ときにはそれ以上に，社会と経済の発展や，個人にとっての社会的，経済的な成功やウェルビーイングの向上，さらに不平等と貧困の削減にとっても鍵となることが，経済学などの社会科学の研究成果により示されてきた（Blanden *et al.*, 2007；Goodman *et al.*, 2015；Gutman and Schoon, 2013；Heckman and Rubinstein, 2001；Heckman *et al.*, 2006；OECD, 2015など）。これを背景に，欧米の教育政策では社会性と情緒の学習（SEL: social and emotional learning）や人格教育（character education）が積極的に議論されたり，導入されたりするようになっている。

もちろんこうした教育は欧米（特に英米）において全く新しいわけではないが，近年の特徴は，教育プログラムの開発や資源配分の根拠としてエビデンスが求められる傾向が強まっていることである。本稿は，まず近年の欧米の学校教育で社会情緒的スキルとその教育が重視されている状況について概観する。続いて，社会情緒的スキルをどのような概念として捉え，どのような観点で測定や分析を行うかについての検討状況をレビューする。日本においても社会情緒的スキル（あるいは，これに分類されるスキル）の教育は決して新しい経験ではないが，国際的な動向や言説の影響を受け，エビデンスに基づく教育政策の文脈で改めて検討を迫られる可能性がある。そこで浮上すると考えられる論点について最後に考察する。

2 欧米の教育における社会情緒的スキルへの着目

欧米の教育といっても一括りにするのは難しい。従来の学校教育の内容の特徴について大きく分けると，英米では教科教育のほか人格教育も重視され，スポーツや文化活動などの課外活動も積極的に行われてきた。他方，ヨーロッパ大陸では教科教育が中心であり，しつけは家庭の責任とされ，スポーツなどは地域のクラブ活動として行われてきた（二宮，2014）。近年のSELの展開も英米がより先進的だといえるが，ヨーロッパ大陸諸国でも程度や方向性の差こそあれ，類似の取り組みが実践されている。近年の展開とその背景について簡単に見ておこう。なお，SELや類似の取り組みは，学校教育だけでなく学校外活動に組み込まれることも多いが，以下では主に政府主導による学校教育への導入の動きを紹介する。

米国では，過去20年ほど教育の成果として学力が過剰に重視されてきた。2002年から2015年にかけて効力をもっていたNCLB（No Child Left Behind）法でも，各州による学力テストの実施や成果目標の設定を通じ，「落ちこぼれ」防止や学力格差是正が目指されていた。こうしたNCLB法に基づく取り組みに対しては，子どもたちの社会情緒的なニーズを軽視しており，経済的に不利な家庭出身者には効果がないばかりか，ときには有害であるとの批判も向けられた（Meier and Woods eds., 2004）。さらに，将来の成功に対する社会情緒的スキルの効果を明らかにしたHeckmanら経済学者の研究が影響力をもち（Heckman and Rubinstein, 2001；Heckman *et al.,* 2006），SELへの期待が高まるようになった。2004年にイリノイ州で初めて就学前教育から後期中等教育におけるSELの基準が開発されて以降，現在では50州全てで就学前教育におけるSELの基準が策定され，多くの州の学校教育にSELが導入されている。連邦政府も予算措置や指針の整備により，エビデンスに基づくSELの実践を推進している（Weissberg *et al.*, 2017）。

英国（イングランド）でも1988年にナショナルカリキュラムとともにナショナルテストが導入され，1990年代にはOfsted（Office for Standards in Education, Children's Services and Skills）による学校監査が導入されたことで，学校教育が学力の達成度により管理される傾向が強まっていた。一方，2003年には政府が『どの子どもも大切（Every Child Matters）』とするアジェンダにおいて，子どもの精神的，身体的な健康や安全を含む，多面的なウェルビーイングを重視する方針を示した。また，2000年代には社会情緒的な側面での教育に関する政府の委託研究報告書（Weare and Gray，2003）や書籍（Weare, 2004）が刊行された。こうした状況下で，学習の社会情緒的側面（SEAL: social and emotional aspects of learning）が注目されるとともに，政府主導で子どもの社会情緒的スキルの向上が推進されるようになる。SEALの枠組みは，ナショナルカリキュラムに2000年に導入されたシティズンシップと人格・社会・健康教育という教科で採用された。2010年から2014年にかけては学校の自律性を高める教育政策のもと，全体としては教育内容に関する政府の主導性が弱まるなかで，数学，英語，科学に関する教育と評価が重要度の高いものと位置づけられた。2015年，これらに続いて重要度の高いものとして，人格と立ち直る力（レジリエンス）の形成が掲げられた（Clarke, 2015；DfE, 2016）。

ヨーロッパ大陸諸国でのSELの導入状況は様々であるが，SELあるいは類似の取り組みが着目，導入されるようになった共通の背景には，1990年代頃から高まっている，メンタルヘルスの維持向上と生涯学習への関心があるようである（Torrent *et al.*, 2017）。前者については，WHO（World Health Organisation）が提起したライフスキル（日常生活で生じる様々な問

題や要求に対して，建設的かつ効果的に対処するために必要な能力）の概念の影響が見られる（WHO, 1994）。後者については，UNESCO（United Nations Educational, Scientific and Cultural Organization）が掲げた学習の4本柱のうち，特に「共に生きることを学ぶ」と「人間として生きることを学ぶ」がSELに関連深いものとみなされる（UNESCO, 1996；Torrent *et al.*, 2017）。

なかでもスペインとフィンランドでは政府の規定するカリキュラムにSELが含まれ，英米と同じように明確な形でSELが推進されている。特にスペインでは英米の影響が強い。スウェーデン，フランス，ポルトガルの学校教育でも社会情緒的スキルは重視されており，シティズンシップ教育などを通じて習得することが期待されている（Torrent *et al.*, 2017）。

3 社会情緒的スキルの概念とエビデンス

このように，欧米の学校教育では近年ますます明確に社会情緒的スキルが重視されるようになっている。そして冒頭でも述べたように，教育プログラムの開発や資源配分の根拠としてエビデンスを求められることが，近年の特徴である。そこで本節は，社会情緒的スキルをどのような概念として捉え，どのような観点で測定や分析を行うかについて，欧米での検討状況をレビューする。ただし，ヨーロッパ大陸諸国では，各国でSELあるいは類似の取り組みについて評価するエビデンスがまだ少ない（Torrent *et al.*, 2017）。そのため英米を拠点とする議論が中心となる。

社会情緒的スキルは，その呼び方が一つに定まっていないだけでなく，それがどのようなスキルを指すかについて説明する定義も様々である。また，スキルというより，心理的，人格的な特性として捉えられる場合もあれば，スキルや特性を活かしながら積極的，効果的に環境に関わろうとする意欲や能力を捉えたコンピテンス（White, 1959）として概念化される場合もある（遠藤，2017）。

米国で1994年に設立され，エビデンスに基づくSELを主導しているCASEL（Collaborative for Academic, Social and Emotional Learning）によれば，SELとは「子どもや成人が，感情を理解して制御し，建設的な目標を立てて達成し，他者に共感してそれを伝え，良好な人間関係を築いて維持し，責任ある意思決定を行うために必要となる知識，態度，スキルを習得し，かつ効果的に応用する過程に関わる」（CASEL, 2012: 4，筆者訳）ものである。この過程で習得され，応用される社会情緒的なコンピテンスあるいはスキルは，自己認識，自己管理，社会認識，関係構築スキル，責任ある意思決定の5つの領域に分類されている。

英国（イングランド）の教育担当省は，社会情緒的スキルを「他の人々と良好な関係を築き，自分自身と自分の感情，思考，行動を理解するスキル」（DfES, 2007: 4，筆者訳）と定義する。こうしたスキルを習得することにより，自発性や楽観性の向上，感情の制御，挫折の克服，自分と他者の権利の認識と擁護，他者との共通性と差異の尊重，持続的な友人関係の構築，他者との協力による問題解決などが期待されるという。また，教育担当省との連携も行っている非営利のシンクタンク，ヤング財団（Young Foundation）のプロジェクトは，教育，仕事，健康などについて長期的に良好なアウトカムをもたらす重要な社会情緒的スキルとして，コミュニケーション，自信と主体性，計画と問題解決，関係構築とリーダーシップ，創造性，回復力と決意，感情の管理の7つのクラスターを導き出している（McNeil *et al.*, 2012）。

これらの定義や分類は，かつては目に見えず捉えどころがないと考えられ，論者により比較的自由に定義されてきた社会情緒的スキルの概

念について，その構成要素を具体的に特定し，測定につなげ，エビデンスに基づいて教育の対象となるか否かを検討しようという意図のもと，体系的に整理したものといえる。教育の対象となり得る重要なスキルとして選定されるどうかは，それが社会的，経済的に意味のあるアウトカムを予測するかどうかにかかってくる。では，意味のあるアウトカムとは何か。そこには学業の達成，就学の継続や進学，問題行動の回避など，子どもにとって現在や近い将来に現れるアウトカムのほか，仕事，収入，健康，社会参加，人間関係の形成などの長期的なアウトカムが含まれる。さらに，そもそも社会情緒的スキルが教育を通じて変容可能（malleable）かが，教育の対象となり得るかどうかを検討する際に問われている。

エビデンスの蓄積と普及により，社会情緒的スキルの重要性についての国際的な認識を一層高めるとともに，各国の教育政策立案に貢献することを目指すOECD（Organisation for Economic Co-operation and Development）は，社会情緒的スキルについて，上述の観点を端的にふまえた定義を提示している。その定義によれば，社会情緒的スキルとは「（a）思考，感情，行動の一貫したパターンに顕在化し，（b）フォーマルあるいはインフォーマルな学習経験により発達し，（c）その人の人生を通じて社会経済的なアウトカムの重要な原動力となり得る個人の能力」（OECD, 2015: 35，筆者訳）である。

社会情緒的スキルが，心理的，人格的な特性や性質として捉えられる場合もあることは既に述べたが，OECD（2015）は，その具体的な構成要素を捉える際に，明示的にパーソナリティ心理学の特性5因子モデル（外向性，共感性，勤勉性，情緒の安定性，開放性），いわゆるビッグ・ファイブを主な枠組みとして援用している。社会情緒的スキルに相当するものを心理社会的スキル（psychosocial skills）と呼んで検討している米国を中心とした研究においても，ビッグ・ファイブの枠組みに依拠した定義が示されている（Burrus and Brenneman, 2017）。そこでは性格とは，「思考，感情，行動に一貫性をもたらす，個人の持続的な内的特性の体系である」（Leary, 2005: 3，筆者訳）と説明される。もちろん社会情緒的スキルを捉えるにあたり，ビッグ・ファイブが唯一の有効な枠組みとされているわけでなく，今後の研究により発展の余地があると考えられている。しかし，ここで注目しておきたいのは，社会情緒的スキルが，決して暗にではなく，明らかに，性格とみなされる要素を含んだ概念となっていることである。

そして社会情緒的スキルが性格とみなされる要素を含むとしても，それが教育の対象となり得るかどうかをエビデンスに基づいて問うことは可能とされている。従来，性格は生まれもった不変のものと考えられることが多かったが，最近の研究成果からは，性格が変容可能であり，しかもSELのような教育的介入による意図的な変容も可能であることが明らかになっている（Burrus and Brenneman, 2017; Walton and Billera, 2017）。また，性格が社会的，経済的に意味のあるアウトカムを予測することを示す研究は古くからある（Bowles *et al.,* 2001; Jencks *et al.,* 1979）。すなわち，スキルを性格と置き換えた場合でも，意味のあるアウトカムの予測可能性という観点と，教育による変容可能性という観点から，教育の対象として検討の余地のあることが示唆されている。

社会情緒的スキルを教育の対象とするといっても，社会情緒的スキルには多様な構成要素があり，具体的にどのようなスキルをどのような教育により向上させるかという問いは残る。まさにこの点もエビデンスをもとに検討すべき課題であるとされ，近年は調査や実験などの方法を用いた研究が進められている。こうした研究成果のレビューについて参照できる文献も多数ある（Berkowitz, 2008 ; Clarke *et al.,*

2015 ; Durlak *et al.* eds., 2017 ; Gutman and Schoon, 2013 ; Lipnevich *et al.,* 2017 ; Weare and Gray, 2003 など）。また，教育の評価研究のレビューを通じて「何が有効か」に関するエビデンスの質を格付けして集約する，米国のWWC情報センター（What Works Clearinghouse）のウェブサイトには，SELに類する教育プログラムの評価結果も掲載されている。

しかし，これまでの研究成果からは，社会情緒的スキルを教育対象とすることが決して単純な課題ではないことも示唆される。社会情緒的スキルの多様な構成要素のうち，あらゆるアウトカムに対してプラスの影響をもたらす，万能な解決策となるようなスキルを特定できているわけではなく，今後の研究でも特定は難しいと考えられている（Gutman and Schoon, 2013）。特定のスキルが，あるアウトカムに対してはプラスの影響をもたらすが，別のアウトカムに対してはマイナスの影響をもたらす場合もある。例えば，ノルウェーのデータに基づく分析では，外向性に関するスキルが抑うつの減少には有効である一方，肥満を増加させるという結果が出ている（OECD, 2015）。また，厳密な意味で因果的効果を明らかにした研究も少ないのが現状である（Gutman and Schoon, 2013; OECD, 2015）。

各国で実施されている個別の教育プログラムの評価だけでなく，全国レベルの調査で児童生徒の社会情緒的スキルについて測定し，場合によっては国際比較を通じ，社会情緒的スキルの育成という観点からそれぞれの国や地域の教育制度の長所や短所を評価し，改善点について模索することも重要な研究課題として指摘されている（OECD, 2015）。米国の全米学力調査（NEAP: National Assessment of Educational Progress）では，2017年からやり抜く力（グリット）や学習意欲に関する質問項目が導入された。国際的な調査であるOECD生徒の学習到達度調査（PISA: Programme for International Student Assessment）では，避けがたい文化的バイアスに対応した調査技術の開発が進められるとともに，2012年から性格的な側面（問題解決への態度など）に関する質問項目が導入された（Bertling *et al.,* 2017）。

4 問われる教育政策の対象

ここまで，近年の欧米でSELを通じた社会情緒的スキルの育成が，エビデンスに基づいた教育政策立案が強調される文脈において重要性を増してきている状況を概観してきた。その際に求められるエビデンスには，本来，様々な性質の知識や根拠が含まれるべきだと考えられる。しかし近年は，ランダム化比較試験の結果を最も高く信頼して教育プログラムの有効性を評価する，「何が有効か」に関するエビデンスが，政策立案において圧倒的に重宝される傾向がある。SELが推進される背景にも，そうした傾向が見て取れる。

「何が有効か」という視点が席巻する状況に対しては，教育の効果に関する技術的な知識のみを不当に強調し，教育に値するものとは何かをめぐる価値の議論を排除したり軽視したりしているという批判もある（Biesta, 2007; Oancea and Pring, 2009）。この批判は，社会情緒的スキルを育成するにはどのような教育が有効かばかりを問う状況に対しても当然向けられるはずである。社会情緒的スキルが教育対象となるか否かについては，それが社会的，経済的に意味のあるアウトカムを予測するかと，教育的介入を通じて変容するかという観点から検討される動向については既に述べた。こうした観点に限っても，エビデンスをもとに明快な結論を導くことは難しい。しかしながら，仮に社会的，経済的な成功に影響を及ぼす社会情緒的スキルと，そのスキルを伸ばす教育プログラムを特定できたとしても，それを学校教育に導入すべきだという結論を導くための論拠が十分に揃ったとはい

えない。あくまで教育の技術的な可能性が明らかになったにすぎず，それを学校教育に導入することが望ましいか否かの価値判断は別に必要である。

特に，社会情緒的スキルには性格と峻別しがたい要素が含まれるため，性格を教育対象とすることの是非について慎重に考えるべき課題がある。性格は能力のように高ければ高いほど望ましいと評価されるものではなく，それぞれの状況や文脈で表れる個々人の特性の違いを意味するものである（遠藤，2017）。その違いを違いのまま尊重することに価値を置くならば，教育的介入を通じて性格を変容させようとすることとそれ自体の是非が根本的に問われるべきだろう。日本ではこれまでも，また教育基本法でも，性格というより人格という用語で，人格形成が教育の目的の一つに掲げられているが，これに対して根本的な疑問を投げかけようというのではない。問われるべきは，特定のアウトカムに向けた生産性の高低に関するエビデンスが提示されたとしても，道徳的には善悪を決めがたいような性格を，本人の意思とは無関係に教育を通じて一定の方向へと変容させようとすることの是非である。

社会情緒的スキルが社会と経済の発展あるいは個人の成功やウェルビーイングに影響し，ひいては不平等や貧困の削減に貢献するというエビデンスが，性格を教育対象とすることに大義を与えるだろうか。ここで，教育を通じた社会情緒的スキルの育成が，ウェルビーイングの向上や不平等と貧困の削減の手段として提案される，唯一の政策オプションではないことを念頭に置く必要がある。変化が求められるのは，個々人の社会情緒的スキルの習得状況というよりは，社会情緒的スキルが使用される環境や文脈，あるいは特定の社会情緒的スキルに報酬が与えられる仕組みのほうにあるかもしれない。また，社会情緒的スキルの育成を妨げる環境（例えば幼少期の貧困や虐待など）の改善も，事後的な教育的介入とは似て非なる政策的対応として重要となり得る。こうした様々な選択肢が考えられるなかで，教育政策にどのような課題を引き受けさせるか検討するには，技術的な可能性を示す「何が有効か」のエビデンスだけではなく，教育的介入の望ましさについての議論も求められる。

5　エビデンスの蓄積を支える議論の必要性

それでは，仮に性格への教育的介入を躊躇する立場に立った場合，性格との峻別が難しい社会情緒的スキルに関するデータを収集することに，全く意義は見出せないだろうか。社会情緒的スキルに限らず，教育分野では成果を定量的に測定することへの抵抗感も小さくなく，「数字の一人歩き」や，定量化しやすいスキルだけが偏重されることへの懸念も多い。たしかにこうした懸念には配慮が求められる。しかし，倫理的に問題がない方法であれば，データを収集，分析し，エビデンスを蓄積することには意義があると考えられる。社会と経済の発展，ウェルビーイングの向上や不平等と貧困の削減が現に重要な課題である以上，その要因の解明に取り組むことは，その要因への教育的介入が可能あるいは妥当かどうかにかかわらず，必要な挑戦だろう。社会情緒的スキルがその要因の一つと目されるなかで，そうした研究からもたらされるエビデンスは，社会情緒的スキルへの教育的介入が妥当かをめぐる民主的な議論を展開するうえでも必要である。

第3節の終わりに述べた，社会情緒的スキルの観点から教育制度を評価することについては，教育プログラムの開発とは異なる意義も見出せる。例えば，次のようなことである。学習を促進する社会情緒的スキルが学齢期の子どもにとって変容可能であるならば，それを不変と捉えて早期から学習環境を階層化することは，家庭

環境が不利で，学校教育により学習意欲を高めることが特に期待される子どもたちにとって逆効果となる可能性もある。こうした問題を明らかにし，階層化の是非を検討したり，階層化による弊害を最小化する制度を設計したりするため，教育制度の評価を目的とした調査で社会情緒的スキルを測定し，エビデンスを蓄積する必要があると考えられている（Bertling *et al.*,2017）。

教育政策の検討に向けてエビデンスから導かれる議論は，それが不十分な現在の日本で想像されているよりも，もっと多角的に展開されるべきだろう。「数字の一人歩き」や特定のスキルの偏重，あるいは教育対象の無用な拡大への懸念は，エビデンスの蓄積を躊躇することによってではなく，エビデンスを解釈し，利用するための枠組みや議論を豊かにすることによって対処すべきことだと考えられる。日本でエビデンスに基づく教育政策の立案を進める際に，欧米に比べて「何が有効か」のエビデンスの蓄積が遅れている実態ばかりが強調されがちである。しかし，社会情緒的スキルの教育と調査をめぐる欧米の動向から，何を教育の対象とするのが望ましいか検討するには，教育の可能性についての技術的なエビデンスの蓄積だけでは不十分であることに気づかされる。スキルであれ，性格であれ，人間の特性をありのままに尊重することと，それを変容させることの間にある葛藤に対処するには，どちらをどのように重視するかをめぐる価値の議論が不可欠である。

文献

Berkowitz, Marvin W., Victor A. Battistich and Melinda C. Bier, 2008, “What Works in Character Education: What is Known and What Needs to be Known”, Larry Nucci and Darcia Narváez (eds.), *Handbook of Moral and Character Education, Routledge*: 414-431.

Bertling, Jonas P., Francesca Borgonovi and Debby E. Almonte, 2017, “Psychosocial Skills in Large-Scale Assessments: Trends, Challenges and Policy Implications”, Anastasiya A. Lipnevich, Franzis Preckel and Richard D. Roberts (eds.), *Psychosocial Skills and School Systems in the 21st Century: Theory, Research, and Practice*, Springer: 347-372.

Biesta, Gert, 2007, “Why ‘What Works’ Won’t Work: Evidence-based Practice and the Democratic Deficit in Educational Research”, *Educational Theory,* 57(1): 1-22.

Blanden, Jo, Paul Gregg and Lindsey Macmillan, 2007, “Accounting for Intergenerational Income Persistence: Noncognitive Skills, Ability and Education”, *The Economic Journal*, 117(519): C43-C60.

Blank, Lindsay, Sue Baxter, Elizabeth Goyder, Louise Guillaume, Anna Wilkinson, Silvia Hummel, Jim Chilcott, Nick Payne, 2008, *Systematic Review of the Effectiveness of Universal Interventions Which Aim to Promote Emotional and Social Wellbeing in Secondary Schools*, University of Sheffield.

Bowles, Samuel, Herbert Gintis and Melissa Osborne, 2001, “The Determinants of Earnings: A Behavioral Approach”, *Journal of Economic Literature*, 39(4): 1137-1176.

Burrus, Jeremy and Meghan Brenneman, 2017, “Psychosocial Skills: Essential Components of Development and Achievement in K-12”, Anastasiya A. Lipnevich, Franzis Preckel and Richard D. Roberts (eds.), *Psychosocial Skills and School Systems in the 21st Century: Theory, Research, and Practice*, Springer: 3-27.

Clarke, Aleisha M., Silvia Morreale, Catherine-Anne Field, Yomna Hussein and Margaret M. Barry, 2015, *What Works in Enhancing Social and Emotional Skills Development during Childhood and Adolescence?: A Review of the Evidence on the Effectiveness of School-based and Out-of-school Programmes in the UK*, WHO Collaborating Centre for Health Promotion Research, National University of

Ireland Galway.

Collaborative for Academic, Social and Emotional Learning (CASEL), 2003, *Safe and Sound: An Educational Leader's Guide to Evidence-Based Social and Emotional Learning (SEL) Programs*, CASEL.

————, 2012, *2013 CASEL Guide: Effective Social and Emotional Learnings Programs - Preschool and Elementary School Edition*, CASEL.

Department for Education (DfE), 2016, *DfE Strategy 2015-2020: World-class Education and Care*, DfE.

Department for Education and Skills (DfES), 2007, *Social and Emotional Aspects of Learning for Secondary Schools (SEAL) Guidance Booklet*, DfES.

Durlak, Joseph A., Celene E. Domitrovich, Roger P. Weissberg, and Thomas P. Gullotta eds., 2017, *Handbook of Social and Emotional Learning: Research and Practice*, The Guilford Press

遠藤利彦 , 2017,「非認知なるものの発達と教育－その可能性と陥穽を探る」国立教育政策研究所編『非認知的（社会情緒的）能力の発達と科学的検討手法についての研究』:15-27.

Goodman, Alissa, Heather Joshi, Bilal Nasim and Claire Tyler, 2015, *Social and Emotional Skills in Childhood and Their Long-term Effects on Adult Life: A Review for the Early Intervention Foundation*, Institute of Education.

Gutman, Leslie Morrison and Ingrid Schoon, 2013, *The Impact of Non-cognitive Skills on Outcomes for Young People: Literature Review*, The Institute of Education .

Heckman, James, J. and Yona Rubinstein, 2001, "The Importance of Noncognitive Skills: Lessons from the GED Testing Program", *American Economic Review*, 91 (2): 145-149.

Heckman, James, J., Jora Stixrud and Sergio Urzua, 2006, "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior", *Journal of Labor Economics*, 24(3): 411-482.

Jencks, Christopher *et al.*, 1979, *Who Gets Ahead? The Determinants of Economic Success in America*, New York: Basic Books.

Leary, Mark, R., 2005, "The Scientific Study of Personality", Valerian J. Derlega, Barbara A. Winstead and Warren H. Jones (eds.), *Personality: Contemporary Theory and Research*, 3rd edition, London : Thomson Learning: 3-26.

Lipnevich, Anastasiya A., Franzis Preckel and Richard D. Roberts eds., 2017, *Psychosocial Skills and School Systems in the 21st Century*: Theory, Research, and Practice, Springer.

McNeil, Bethia, Neil Reeder and Julia Rich, 2012, *A Framework of Outcomes for Young People*, Young Foundation.

Meier, Deborah and George Woods eds, 2004, *Many Children Left Behind: How the No Child Left Behind Act is Damaging Our Children and Our Schools*, Boston : Beacon Press.

二宮晧 , 2014,「グローバル化する世界の学校を旅する」,二宮晧編『新版 世界の学校 ― 教育制度から日常の学校風景まで』学事出版:8-13.

Oansea, Alis and Richard Pring, 2009, "The Importance of Being Thorough: On Systematic Accumulations of 'What Works' in Education Research", David Bridges, Paul Smeyers and Richard Smith (eds.), *Evidence-based Education Policy: What Evidence? What Basis? Whose Policy?*, Malden: Wiley-Blackwell:11-35.

Organisation for Economic Co-operation and Development (OECD), 2015, *Skills for Social Progress: The Power of Social and Emotional Skills*, Paris : OECD Publishing.

Torrente, Catalina, Anjali Alimchandani and J. Lawrence Aber, 2017, "International Perspective on SEL", Joseph A. Durlak, Celene E. Domitrovich, Roger P. Weissberg, and Thomas P. Gullotta (eds.), *Handbook of Social and Emotional Learning: Research and Practice*, New York: The Guilford Press:566-587.

United Nations Educational, Scientific and Cultural Organization (UNESCO), 1996, *Learning: The Treasure Within*, UNESCO.

Walton, Kate, E. and Kimberly A. Billera, 2017, "Personality Development During the School-Aged Years: Implications for Theory, Research, and Practice", Anastasiya A. Lipnevich, Franzis Preckel and Richard D. Roberts eds., *Psychosocial Skills and School Systems in the 21st Century: Theory, Research, and Practice*, Springer: 93-111.

White, Robert W., 1959, “Motivation Reconsidered: The Concept of Competence”, *Psychological Review*, 66 (5):297–333.

Weissberg, Roger P. Joseph A. Durlak, Celene E. Domitrovich and Thomas P. Gullotta, 2017, “Social and Emotional Learning: Past, Present, and Future”, Joseph A. Durlak, Celene E. Domitrovich, Roger P. Weissberg, and Thomas P. Gullotta (eds.), *Handbook of Social and Emotional Learning: Research and Practice*, New York: The Guilford Press: 3-19.

Weare, Katherine, 2004, *Developing the Emotionally Literate School*, Los Angeles: Sage Publishing.

Weare, Katherine and Gay Gray, 2003, *What Works in Developing Children's Emotional and Social Competence and Wellbeing?*, Nottingham : Department for Education and Skills.

World Health Organisation (WHO), 1994, *Life Skills Education for Children and Adolescents in Schools: Introduction and Guidelines to Facilitate the Development and Implementation of Life Skills Programmes*, WHO.

特集論文 4

文部科学省の全国学力・学習状況調査の意義と問題点

川口俊明
福岡教育大学教育学部 准教授

1 全国学力・学習状況調査の概要

本稿では，文部科学省が2007年から実施している全国学力・学習状況調査（以下，全国学力調査と呼ぶ）について，社会調査の視点から見た意義と問題点について論じる[1]。

はじめに，公にされている全国学力テストの目的や概要を確認しておこう。文部科学省が示している全国学力調査の目的[2]を整理すると，全国学力調査には大きく二つの役割が期待されている。一つは，学力調査によって教育施策を検討するために必要な基礎資料を得るという点である。子どもたちの学力は，教育政策を評価する指標となる。行政の説明責任が強調される昨今，教育政策の成果と課題を検証するために，「教育政策のための学力調査」が必要とされているという訳である。もう一つが，学力調査の結果を一人ひとりの児童生徒の指導の改善・充実に役立てるという点である。すべての学校が一人ひとりの子どもの実態を把握し，指導の充実・改善を行うためには，毎年度すべての子どもが学力調査に参加しなければならない。この「指導のための学力調査」という発想は，調査が実施されて10年目を迎えた折に，全国的な学力調査に関する専門家会議から出された「全国的な学力調査の今後の改善方策について（まとめ）」[3]の冒頭でとくに強調された点であり，全国学力調査を毎年度，悉皆実施する根拠となっている。

つまり，現行の全国学力調査は，「教育政策のための学力調査」と「指導のための学力調査」という異なる二つの目的を一つの調査の中で同時に達成しようとしているわけである。その試みが成功しているか否かについては後で述べるとして，続いて全国学力調査の調査対象について確認しておこう。

全国学力調査の対象学年は，小学6年生，および中学3年生である。調査教科は，主として国語と算数（数学）であり，テストはペーパーテスト形式で実施されている。また，テストに加えて，児童生徒の生活実態に関する調査が行われる他，学校の管理職を対象にした学校質問紙調査も行われている。すでに述べたように，すべての児童生徒の指導に活かすという観点から，調査は悉皆実施である（ただし，民主党政権時の2010年，2012年は抽出実施）。

調査結果の公表には，二つの段階がある。一つは，調査問題の全文・回答例・出題の意図や狙いを示した解説資料であり，これらは調査実施直後の4月下旬に公開される[4]。もう一つは，8月頃に開示される調査結果であり，こちらは各都道府県の平均点を始め，テストの得点と関連のある項目を示した分析等が公開される。なお，調査の個票は，2018年時点では非公開とされており，研究者が自由に二次分析に利用する

ことはできない。

ここまでに述べてきた，毎年実施される悉皆の全国学力調査は「本体調査」と呼ばれている。この「本体調査」に加えて，平成25年以降，数年おきに「きめ細かい調査（経年変化分析調査，保護者に対する調査，教育委員会に対する調査）」が抽出で実施されている。また，文部科学省の委託事業として，調査で得られたデータを元にした追加分析調査も実施されており，全国学力調査で得られたデータを元にした分析や小規模の調査を付け加えた分析が行われている。本稿では，こうした「きめ細かい調査」や追加分析調査についても論じることにする[5)]。

2 全国学力・学習状況調査の意義

社会調査の観点から見た全国学力調査の意義は，この調査が，日本の学力実態や教育政策の効果を明らかにしてきたという点にある。アメリカのように二次分析に利用可能な学力調査を豊富に有している国もあるが，日本はそのような学力調査に乏しい。現在でも，日本で行われる学力研究は，一部の地域から抽出したデータに基づいたものや，特定の自治体と協力して進められているものがほとんどである。ナショナルサンプルかつ二次分析が可能なデータとなると，国際学力調査であるPISAやTIMSS，あるいは慶應義塾大学パネルデータ設計・解析センターが実施する「日本子どもパネル調査」しかないのが現状である。このような状況において，日本全体を把握できる全国学力調査の存在価値は大きい。以下では，全国学力調査を利用した研究の成果を，いくつかの観点からまとめた。

一点目は，保護者の学歴や年収に代表される児童生徒の社会経済的地位（SES: Socio-Economic Status）と学力の関連に関する研究である。「子どもの貧困」「学力格差」が社会問題となって久しいが，日本の小中学校で保護者の学歴や年収といった項目を尋ねることは，調査に対する忌避感や個人情報の保護もあって，容易なことではない。そのため，学力格差に関心を持つ研究者たちは，子どもに尋ねた情報（家庭にある本の冊数等）をSESの代替指標にするか，協力を得られた一部の地域や学校で詳細な調査を実施するといった方法を使って研究を進めてきた（苅谷・志水編，2004；耳塚，2007 等）。

その点，2013年度の全国学力調査で実施された「きめ細かい調査（保護者に対する調査）」は，日本全体の小中学校から抽出したデータを用い，学力格差の実態を明らかにした「画期的な」試みであったと言うことができる。当該データは，学力格差の実態解明のみならず，格差を克服する学校の分析など，これまで一部の地域のデータを基に行われていた研究を，ナショナルサンプルで検証することを可能にした（お茶の水女子大学，2014）。「きめ細かい調査」のデータは，その後，卯月・末冨（2015）による子どもの貧困と学力に関する研究でも利用され，貴重な成果を生み出している。

二点目は，教育経済学に代表される，学級規模や教育政策の効果に関する研究である。これらの研究は，主として学級規模が学力に与える影響に着目したものが多い。そこには，学校単位の集計データを利用したAkabayashi & Nakamura（2014），単学級データを利用した妹尾他（2013），個票データを利用した妹尾他（2014），都道府県の平均点を利用した山本・井上（2015），そして先に述べた「きめ細かい調査」のデータを使った妹尾他（2016）等がある。学級規模以外にも，都道府県別の平均点を利用し，教員の質が学習に与える影響等を推定した二木（2017）や，学校質問紙を利用し習熟度・少人数指導の効果を検討した山森・奥田（2014）がある。

三点目として，学力の経年変化に関する研究がある。全国学力調査は，年度ごとにテスト項

目がすべて刷新されるため，異なる年度間の成績を比較することができない。この点を補完するために行われているのが「きめ細かい調査（経年変化分析調査）」である。同調査に関連して，項目反応理論や重複テスト分冊法といったPISA・TIMSSで利用されている技術を適用していくための調査研究が東北大学によって実施され，全国学力調査に近年のテスト理論の技術を活かす試みが行われている（東北大学，2011；2012；2013；2014）。

ここまで全国学力調査を利用した研究を三つの領域に分類したが，そこには大別して，文部科学省等に許諾を得て非公開になっている個票データを利用する研究と，すでに公開されている都道府県の平均点等の情報を利用する研究の2パターンがある。現時点では全国学力調査の個票が公開されていないため，ほとんどの研究が前者である。この場合，データを利用できるのは特定の組織に所属する研究者に限られるし，分析結果を第三者が検証することも難しい。全国学力調査の個票の公開について現在議論が行われているが，「本体調査」のデータはもちろん，「きめ細かい調査」や追加分析調査のデータを含めて，できるだけ多くの研究者がデータを活用できる制度づくりが求められている。

3　全国学力調査の問題点は何か

ここまで全国学力調査の意義を述べてきた。しかし，全国学力調査は，社会調査の観点からみて深刻な欠点を抱えている。それは，調査に異なる二つの目的が混在しているという点である。冒頭に指摘したように，全国学力調査は，「教育政策のための学力調査」と「指導のための学力調査」という二つの目的を持っている。近年強調されているのは後者の側面であり，毎年度，調査を悉皆実施する根拠となっている。悉皆実施をしたデータを集計すれば，「教育政策のための学力調査」として利用できると考えているからであろう。しかし，毎年度の学力調査の悉皆実施には，以下に述べるようなさまざまな問題があり，そのデータを教育政策に利用するのは容易ではない。

ここでは，すでに行われている全国学力調査への批判・検討（木村，2006；荒井・倉本，2008；苅谷，2008；北野，2016；中澤，2016等）を踏まえつつ，全国学力調査を毎年度，悉皆実施することに伴う問題を三点ほど指摘しておきたい。

一点目は，テストの質が確保できないという点である。現行の全国学力調査のように，毎年度調査を実施し，しかもテスト項目は調査終了直後に公開するという調査形式では，「教育政策のための学力調査」に必要なテストの精度を確保する時間が足りない。学力調査の設計は，それほど簡単な作業ではない。テストが測定する能力を設定し，その能力を把握するために必要と考えられるテスト項目を大量に作成する必要がある。作成したテスト項目の質は，予備調査を行って明らかにしなければならないし，テスト実施後も，個々のテスト項目が想定通りに機能したかどうか確認する必要がある。PISAやTIMSSでも，それぞれの調査はテスト項目の作成と検証に年単位の時間をかけている。現在のように，毎年，テスト項目をすべて一から作り直すという形式では，よほど人員をテストの作成に配置しないかぎりテストの質を担保できない。

さらに全国学力調査は，すべての受験者が同一のテストを受験することを想定している。受験者の集中力を考えれば，一定時間内に回答できるテスト項目の数には上限がある。そのため，全受験者が同一のテストを受験するという現在の形態では，幅広いテスト項目を出題することができず，テストの出題範囲は偏らざるを得ない。このような偏りをもったテスト結果を基に，教育政策のあり方を検討するのは適切とは言え

ない。テスト項目の幅を広げるという観点からは，個々人に一部のテスト項目が重複した異なるテストブックレットを与える重複テスト分冊法を導入することも考えられるが，この場合，異なるブックレット間をつなぐテスト項目は厳重に秘匿する必要がある。この場合，秘匿されたテスト項目は指導のために利用することができなくなるので，「指導のための学力調査」という目的と矛盾する可能性が高い。

悉皆実施に伴う問題の二点目は，学力調査を教育政策に活かせる形で分析するための情報が不足しがちになるという点である。一時点の学力調査から得られるデータでは，相当に分析をくふうしないかぎり，相関関係と因果関係を区別できず，教育の影響を把握することは難しい。そのため，昨今の社会科学では，同一個人を追跡したパネル調査の分析が重視されるようになっている。さらに，SESに代表される児童生徒の家庭環境が成績に影響を与えていることはよく知られた事実であり，学力格差を縮小するといった政策立案のためには，保護者の情報が欠かせない。つまり，教育政策の立案に必要な情報を得るためには，現在の全国学力調査では不十分であり，小学6年生・中学3年生以外の学年での調査や，SESを把握するための保護者調査の実施が必要となる。しかし，毎年度の悉皆実施を前提とするかぎり，こうした追加の調査を行うことは，予算的・時間的な制約を考えると現実的とは言えない[6]。

三点目として，個々人の成績を評価する学力調査の悉皆実施は，各学校・教員に与える影響が大きく，どうしてもハイ・ステークスなテスト（≒結果が被評価者に強い影響を与えるテスト）になりやすいという点が挙げられる。全国学力調査で得られた各学校の平均点を学校評価・教員評価に利用しようと考える人は少なくないから，学校・教員側には，無理をしてでも学力テストの得点を向上させようという動機が生まれる。こうした状況は，不正を誘発しやすいし，そこまで行かなくても，テストに極端に偏った指導が蔓延する可能性がある。これでは調査結果に統制不能なバイアスが生まれ，教育政策の基礎資料とすることができない。悉皆実施をしなければ，こうしたリスクは減らせるはずである。

ただ，三点目に関連して気になるのは，全国学力調査には，そもそも学力調査それ自体を学校現場へのメッセージとして使おうとしている節が見受けられるという点である。全国学力調査の出題範囲・内容に関する指針には「調査を実施することによって，各教育委員会や各学校に対して，学習指導要領に示される内容等を正しく理解するよう促すとともに重視される力を子どもたちに身に付けさせるといった国としての具体的なメッセージを示す」という文言が存在する[7]。全国学力調査を利用して，教育委員会や学校の指導を変えようというわけである[8]。通常，調査は実態を把握するために行うものであり，対象に直接影響を与えることを狙うものではない。その意味では，全国学力調査は「調査」と銘打っているものの，社会調査の範囲を逸脱している。

4 なぜ,全国学力調査の問題点は放置され続けるのか

以上のような全国学力調査の問題点は，本稿が初めて指摘したわけではない。悉皆実施に伴う問題点は，全国学力調査が実施される以前から指摘されている（苅谷，2008：138）。それでは，なぜ全国学力調査の問題は放置されるのか。ここでは，次の二点を指摘しておきたい。

一つは，全国学力調査に関わる人々に，「政策のための学力調査」を運用するために必要な最低限のリテラシーが不足しており，しかも，それが改善しない構造があるという点である。日本では，教育に関心を持つ政治家や一般の市民はもちろんのこと，全国学力調査に直接関わ

る委員ですら，全員が望ましいテストのあり方に関する知識を持っているわけではない。項目反応理論をはじめとしたテスト理論の知識が必要とされる昨今の大規模学力調査は，専門知識が無ければ，調査設計はもちろん，その結果を理解することさえ難しい。しかし木村（2006）が指摘するように，全国学力調査の設計は，必ずしもテスト設計を専門としない人々によって議論されていた。

運用する側にリテラシーが不足しているという点では，日本の教育行政の雇用形態の影響も無視できない。組織の一員としてあらゆる仕事をこなすことが求められる，いわゆる「メンバーシップ型（濱口，2009）」の雇用形態を採用する日本の教育行政では，職員が数年で他部署に異動し，新しい仕事をする中で仕事を覚えていくという働き方が主流である。これでは，テストの専門家を学力調査の担当に配置することも，育てることも難しいだろう。

さらに，木村（2010）は，学力調査に関わる機会の多い学校の教員を養成する教員養成系においても，学力調査に関わるリテラシーを学ぶ機会が限定されていることを指摘している。全国学力調査の課題を肌で感じているのは，おそらく学校現場の教員である。しかし，かれらも望ましい調査のあり方について知る機会が少ないのだから，現行の全国学力調査のあり方を批判する声が高まらないのも道理であろう。

以上のような，構造的な問題を含む関係者のリテラシー不足に加えて，ここでもう一つ指摘しておきたいのは，日本の多くの教員が考えるテストと，「政策のための学力調査」に求められる要素が，大きく乖離しているという点である。日本の小中学校の授業では，学習指導要領の「定着」が目指されている。そのため，学校の教員にとってのテストとは，テストの専門家が考えるような児童生徒の能力を測るものではなく，個々のテスト項目の正誤を確認し，間違えた箇所を修正するための資料（≒「指導のためのテスト」）である。こうした考えが誤っている訳ではない。指導の補助としてテストを使い，学習指導要領の「定着（≒全問正解の100点）」を狙う姿勢は，そのテストが教室の中で行われるものであるかぎりは，好ましいとすら言える。

しかし現状，「指導のためのテスト」という考えは，教室を飛び越え，「指導のための学力調査」として，全国学力調査に適用されてしまっている。調査終了直後にすべての問題を開示し，その出題の意図を示したり，調査結果に指導のポイントを例示したりといった全国学力調査の特徴がそれである。全国学力調査が小学6年生や中学3年生を対象にするのも，小学校6年間（あるいは中学校3年間）の学習指導要領の「定着」を確認するためだろう。「指導のための学力調査」という文部科学省が掲げた全国学力調査の目的は，日本の学校教育におけるテストのあり方と親和性が高く，学校・教員が受け入れやすいものになっているのである。

全国学力調査が悉皆実施されるまでの過程に着目した鳶島も，「すべての子どもの学力保障」というロジックが，学力調査の悉皆実施を正当化していく過程を描いている（鳶島，2010）。2007年に全国学力調査が開始されて以降，一部で批判はあるものの，大規模な反対運動が起こっていない現状を見るかぎり，「指導のための学力調査」というロジックは，一定の説得力をもって，学校や社会に受け止められているようである。

以上のことを踏まえれば，10年を超えて全国学力調査が継続されている現状は，「政策のための学力調査」を扱うことのできる組織や文化が，日本には根付いていないことを示していると言える[9]。結局のところ日本では，テスト理論や社会調査の知見が，教育政策の立案にとって不可欠の要件とは見なされていないのである。

5 今後のゆくえ

このような状況は，今後変わっていくのだろうか。関係者のリテラシー不足や人事異動という構造上の問題，そして，学習指導要領の「定着」を目指す学校と「指導のための学力調査」の親和性を考えると，残念ながら，この状況がしばらく続くことが予想される。

ただ，一部に変化の兆しも見える。その一つが，教育経済学者を中心とした学力調査の専門家たちが自治体と協力して実施している「能力を測る」学力調査の出現である。2016年から埼玉県教育委員会が実施している埼玉県学力・学習状況調査（以下，埼玉県学力調査）[10]は，児童生徒の「学力の伸び」を把握するために，項目反応理論をベースに設計され，テスト項目も非公開で実施されている（石川他，2017）。その成果もすでに出始めており，外山他（2017）のような興味深い知見も得られている。こうした研究成果が，テスト理論の有用性を示すことで，現在の全国学力調査が変わっていく可能性はある。

ただし，筆者はそこに若干の懸念も覚えている。それは，埼玉県学力調査が「政策のための学力調査」と親和性が高いとは言えない学力調査の悉皆実施を維持している点である。児童生徒の「伸び」を測定し，指導や評価につなげるためには，悉皆実施でなければならないのかもしれない。しかしそれは，学力調査の結果を学校評価・教員評価と結びつけられるということに他ならない。これは，現行の全国学力調査以上にハイ・ステークスな学力調査が生まれる可能性があることを意味している。社会科学の知見に裏打ちされた「科学的な」学力テストが学校を統制してしまう危険性については，アメリカの事例がよく知られている（Apple, 2006）。その意味では，埼玉県学力調査が学校に与える影響について，今後，慎重に見極めなければならない。

とはいえ，現行の全国学力調査を維持することは，社会調査の観点からは好ましくない。このままでは毎年数十億円の税金が，教育政策のためには活用しがたい調査に浪費されることになる。全国学力調査を「教育政策のための学力調査」として利用したいのであれば，今なすべきは，個票の公開を前提に，保護者調査・教員調査を含む抽出のパネル調査として設計し直すことである。個々の学校や教育委員会が，自らの政策立案や指導のためにより細かいデータを必要とするのであれば，それはかれらが自身の裁量で行えばよい。わざわざ国が「教育政策のための学力調査」の精度を下げてまで，「指導のための学力調査」を行う必要は無い。

残念ながら，今のところ教育測定に関する知見は，日本では，日本行動計量学会や日本テスト学会など，一部の専門家の間で議論されるものに留まっており，その他の教育学関連領域で一般的に利用される段階には至っていない。この状況を変えるためには，教育に関わる人々が，もう少しテストに関わる知識を学ばなければならないだろう。全国的な学力調査を「何のために」，そして「どのように」行うべきなのか。全国学力調査が開始されて10年以上が経過した今，私たちは，この問題を真剣に論じる時期にきている。

注

1) 全国学力調査に関しては，その競争主義的な性格（志水，2009）や，教育の国家統制（中嶋，2008）が論じられることもあるが，本稿では社会調査の側面に限定して論じる。

2) 文部科学省ホームページ「全国的な学力調査（全国学力・学習状況調査等）」http://www.mext.go.jp/a_menu/shotou/gakuryoku-chousa/zenkoku/1344101.htm

3) 全国的な学力調査に関する専門家会議「全国的な学力調査の今後の改善方策について（まとめ）」http://www.mext.go.jp/a_menu/shotou/gakuryoku-chousa/__icsFiles/afieldfile/2017/03/30/1383338_1.pdf

4)「全国学力・学習状況調査」は毎年4月の第3もしくは第4火曜日に実施される。

5) これまでの調査結果については，国立教育政策研究所のウェブサイトを参照のこと。http://www.nier.go.jp/kaihatsu/zenkokugakuryoku.html

6)「教育政策のための学力調査」という観点からすると，教員調査が含まれていない点も問題である（苅谷，2008：161）。

7)「全国的な学力調査の具体的な実施方法等について（報告）」のうち，4. 調査問題及び質問紙調査について（http://www.mext.go.jp/b_menu/shingi/chousa/shotou/031/toushin/attach/1397245.htm）を参照のこと。

8) 2010年の専門家会議の議事要旨（第2回）にも，専門家会議の委員の発言として「悉皆調査による支援をしないと，個々の先生は関心を持たない」「抽出に変わり，調査に関係ない学校は，雰囲気がだれている」といった意見が記載されている。http://www.mext.go.jp/b_menu/shingi/chousa/shotou/074/gijigaiyou/1299002.htm

9) こうした状況は，テストの専門家のあいだで「日本のテスト文化」の問題として論じられている。日本のテスト文化については，たとえば荒井・前川（2005）や柴山（2008）を参照のこと。

10)「埼玉県学力・学習状況調査」https://www.pref.saitama.lg.jp/f2214/gakutyou/20150605.html

文献

荒井清佳・前川眞一，2005，「日本の公的な大規模試験に見られる特徴」『日本テスト学会誌』1：81-92。

Akabayashi, H. and Nakamura, R., 2014, "Can Small Class Policy Close the Gap? An Empirical Analysis of Class Size Effects in Japan", Japanese Economic Review, 65: 253-281.

Apple, M. W., 2006, Educating the "Right Way", Routledge（太田直子訳，2008，『右派の／正しい教育』世織書房.）

荒井克弘・倉元直樹，2008，「全国学力調査への期待」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：229-240。

二木美苗，2017，「子どもの学習に対する教員の質の効果」『日本経済研究』74：56-83。

濱口桂一郎，2009，『新しい労働社会』岩波新書。

石川善樹・伊藤寛武・植村理・田端紳・外山理沙子・中室牧子・分寺杏介・星野崇宏・松岡亮二・山口一大，2017，「子どもの能力を計測するための学力テストの現在と展望」RIETI Policy Discussion Paper Series 17-P-010。

苅谷剛彦，2008，『教育再生の迷走』筑摩書房。

苅谷剛彦・志水宏吉編著，2004，『学力の社会学』岩波新書。

木村拓也，2006，「戦後日本において『テストの専門家』とは一体誰であったのか？」『教育情報学研究』4：67-99。

————，2010，「日本における『テストの専門家』を巡る人材養成状況の量的把握」『日本テスト学会誌』6：29-49。

北野秋男，2016，「わが国の学力調査体制の実態と課題」『教育學雑誌』52：1-14。

外山理沙子・伊藤寛武・田端紳・石川善樹・中室牧子，2017，「負のピア効果」RIETI Discussion Paper Series 17-J-024。

耳塚寛明，2007，「小学校学力格差に挑む だれが学力を獲得するのか」『教育社会学研究』80：23-39。

中嶋哲彦，2008，「全国学力テストによる義務教育の国家統制」『教育学研究』75（2）：157-168。

中澤渉，2016，「教育政策とエビデンス」志水宏吉編『社会のなかの教育』岩波書店：73-101。

お茶の水女子大学，2014，『平成25年度全国学力・学習状況調査（きめ細かい調査）の結果を活用した学力に影響を与える要因分析に関する調査研究』。

妹尾渉・北條雅一，2016，「学級規模の縮小は中学生の学力を向上させるのか」『国立教育政策研究所紀要』145：119-128。

妹尾渉・北條雅一・篠崎武久・佐野晋平，2014，「回帰分断デザインによる学級規模効果の推定」『国立教育政策研究所紀要』143：89-101。

妹尾渉・篠崎武久・北條雅一，2013，「単学級サンプルを利用した学級規模効果の推定」『国立教育政策研究所紀要』142：161-173。

柴山直, 2008,「日本のテスト文化について」『人事試験研究』No.208：2-13。

志水宏吉, 2009,『全国学力テスト』岩波ブックレット。

鳶島修治, 2010,「全国学力テストの悉皆実施はいかに正当化されたか」『社会学年報』39：75-86。

東北大学, 2011,『全国規模の学力調査における重複テスト分冊法適用の試み』。

———, 2012,『全国規模の学力調査における重複テスト分冊法の展開可能性』。

———, 2013,『全国規模の学力調査におけるマトリックス・サンプリングにもとづく集団統計量の推定について』。

———, 2014,『東日本大震災の学力への影響』。

卯月由佳・末冨 芳, 2015,「子どもの貧困と学力・学習状況」『国立教育政策研究所紀要』144：125-140。

山森光陽・奥田麻衣, 2014,「児童生徒－教師比の縮減を目的とした追加的教員配置の有無による小学校算数学力調査正答率の学校平均の比較」『国立教育政策研究所紀要』143：197-207。

山本信一・井上麻央, 2015,「学級規模と学力」『生活経済学研究』41：55-64。

特集論文

5

「エビデンスに基づく教育」への質的研究の立場からの批判的検討
課題と可能性

酒井 朗
上智大学総合人間科学部 教授

1 エビデンスに基づく教育政策の推進

教育の制度やカリキュラムを改革する上で，科学的根拠，すなわちエビデンスに基づく議論が必要だという指摘がなされている。その背景には，教育や知識の重要性に対する認識の高まりと教育費支出における説明責任の履行への要請がある。そして，そうした社会的要請に対処するために教育研究の質や有効性を向上させなければならないという課題意識がある（岩崎，2011）。これらの議論は相互に重なり合い，教育分野において「エビデンスに基づく教育」という用語が広く浸透している（玉井・藤田，2017）。本稿では，教育に関する質的研究に従事する研究者の議論の紹介を通じて，こうした社会の動向が教育研究や教育現場にもたらす問題を批判的に読み解くとともに，教育現場へのエビデンスの提供に関して，質的研究にどのような可能性があるのかについて検討する。

岩崎（2011）が指摘しているように，エビデンスという言葉は「研究者にとっては，教育科学の発想とともに，教育統計や教育測定などのように，技術化することで科学的客観性を担保しようとする志向を象徴するもの（102頁）」であり，政策立案の根拠として通常期待されるエビデンスは量的研究によるものが多い。このため，質的研究の立場から現在の動向を論評しようとする場合，国内には議論の手がかりになる資料は少ない。だが，国外の研究動向に目を転じてみると，イギリスで20年前にこの問題をめぐる有名な論争が生じていたことが分かる（Oancea, A. & Pring, R., 2009, 訳，2013）。

論争のきっかけは，イギリスの著名な教育社会学者であるD.H.ハーグリーブスが，「研究に基づく専門職としての教職―可能性と期待」という演題で，1996年に同国の教員養成研修局（Teacher Training Agency）で記念講演を行ったことにある（Hargreaves, 1996）。この講演において彼はそれまでの教育研究を批判し，エビデンスとして活用される教育研究を推進していくべきだとの論を展開し，その後のイギリス教育政策に多大な影響を与えた。石井（2015）も指摘しているように，この講演を契機にして「エビデンスに基づく教育」の是非が問われ，論争が繰り広げられた。そして，その過程で質的研究の立場からも，技術的合理性や量的研究重視の傾向の問い直しや，教育研究が教育現場に提供し得るエビデンスがどのようなものかという議論が展開されたのであった。

そこで本稿では，この議論に加わった代表的な2人の質的研究者の批判や提案を取り上げる。最初に紹介するのは，教室のエスノグラフィーに携わり，質的方法論についての著作も多いマーティン・ハマースレーである（Hammersley, 1989; Hammersley & Atkinson, 1983）。彼は，教育の現場や政策に教育研究がエビデンスとして活用されることは，研究にとっ

ても教育実践にとっても重大な問題が生じると批判した（Hammersley, 1997）。彼の批判は，教師の専門職性，教育という行為の本質に関わる論点が含まれており，単なる方法論的な批判にとどまるものではなかった。このことから分かるのは，方法論の議論は，教育という営みや教師の役割をめぐる議論とセットであり，質的方法からの検討とはそうした教育論を含めた形での議論とならざるを得ないということである。

もう1人は，質的研究によるアクションリサーチの提唱で有名なジョン・エリオットである[1)]。彼は，まず教育研究が貢献すべき教育の目標を丁寧に検討した上で，その目標の実現に最も効果的な手立てとして，事例研究によるエビデンスの収集が望ましいとした（Elliott, 2001）。

最終節では以上の議論をもとに，「エビデンスに基づく教育」の実施が強く求められる現在の動向を，教育研究はどのように受け止め，対応していくべきかについて考察する。

2 D.H. ハーグリーブスによる教育研究の変革の訴え
エビデンスの希求

D.H. ハーグリーブスは，イギリスの中等学校の生徒文化に関するエスノグラフィーを1960年代末に著して世界的に知られるようになった教育社会学者であり，我が国の教育研究にも多大な影響を及ぼした（Hargreaves, 1967）。その後，彼は同国のナショナル・カリキュラムの作成を担う「資格・カリキュラム機構」の機構長を務めるなど，教育行政に深くかかわった。こうした経歴を持つハーグリーブスは，1996年の教員養成研修局での講演で，医師と教師の専門性を比較し，社会的評価の高い医師の専門性は医学研究の膨大な蓄積に支えられているのに対し，教師の専門性は，その基盤となるべき教育研究が積み上げ的なものになり得ていないがゆえに低いものにとどまっていると批判した。

講演の中でハーグリーブスは，「パラダイムシフトと言われるような研究上の流行の変化により，研究の流れが途切れてしまっている。社会科学は積み上げ的な学問だというのは幻想にすぎないというポストモダン論者の批判が，教育研究の意義を貶めている」と批判した。その上で彼は，教育専門職は教育研究を必要としているが，研究が実践に貢献しようとするのであれば，その研究は現状とは全く違ったものにならなければならないと述べた。

ハーグリーブスが変革を求めたのは，応用研究の分野であった。医学における応用研究は，病院での医師による日常的な医療実践に基づいたものであり，医療はその中で見出されたエビデンスに基づいて遂行される。最も重要な研究は，医師が行う治療効果の測定に関する研究である。エビデンスに基づいた医療は，医療行為の改善に直接的かつ即時的な効果を有しており，患者の利益になるため，医師は知識をアップデートしようと躍起になる。

これに対して教育研究には，教師の実践の効果や影響に関する十分なエビデンスの蓄積が乏しい上，教師自身もエビデンスに基づいた研究がないことに問題を感じていない。教師は今でも自分の好みに基づいて実践を進めなければならず，自身の経験の中での試行錯誤から学ぶことが多い。ハーブリーブスはこのように指摘した。さらに講演の後半では，現状では教育効果の高い実践に関する明確なエビデンスは無いが，「指導方法の変更が学習に及ぼす効果」という基本課題について問いを発しなければならないとも述べた。

3 ハマースレーによる批判

ハマースレーは，このハーグリーブスの講演にすぐに反応し，翌年の1997年，イギリスの教育研究雑誌に批判論文を掲載した。その論点の１つは，教育研究が積み上げ的に進展していな

いと批判するハーグリーブスへの反批判である。要点は以下の通りである。

①　研究上の積み上げとはエビデンスの蓄積だけではない。多くの教育研究者は理論的な発展を担ってきた。

②　20世紀の3分の2（1960年代半ば：筆者注）までの教育研究は，ハーグリーブスが推奨する，教育効果の検証，すなわち「科学的なモデル」に傾注してきたのであり，この20年余りの間にそうした「実証主義的」研究から質的研究に移行してきたのは，教育効果研究が有効な知見を産出しえなかったからである。

③　1970年代に教育研究が質的研究へと移行したのは，「実証主義的」研究が解明しえない問題に対する批判が巻き起こったからでもあり，ハーグリーブスは，その中で主導的な役割を果たした研究者の1人であった。

④　批判の急先鋒はWinch（1958, 訳, 1977）らによる哲学的議論，すなわち，人間の社会生活は自然科学者が扱う物理的世界とは性質が全く異なるものであるという批判であり，社会科学や教育研究者には自然科学が生み出すような知の産出はなしえないと結論付けられた。

⑤　社会現象においては，因果関係に関して決定的な知見を得るには相当難しい問題があることは否定しえない。このことが理由の１つとなって，教育研究者は，他の社会科学者同様に，哲学的論争や方法論をめぐる論争に巻き込まれた。

ハーグリーブスの講演に対するハマースレーの批判の第二のポイントは，授業という営みの性質の理解と，そこへの研究の貢献のあり方に関してである。その要点も同様に列記すると以下の通りである。

①　ハーグリーブスは研究の実践への貢献に関してきわめて狭い見方に立っている。研究は実践上の課題に一番有効な指導法は何かを教師に明示するものでなければならないと考えられている。しかし，教師の直面する課題がいかなる問題であるか，そのこと自体が研究により解明されるべきである。つまり，その課題がテクニカルな問題であるのかどうかを問わなければならない。

②　教えるという営みに関する哲学や心理学，社会学の近年の研究によれば，実践的であることの本質は技術ではない，つまり，教えるというのは，あるルールに沿って進めるものではなく，実践者がその場で判断をしながら進める行為である。研究的な知見に基づいて授業がなされうるという考え方自体に疑問符が付くのであって，授業とは経験や見識，その場固有の知識，判断に依存するものである。

③　授業のねらいは多様であり，それ自体が議論の対象である。教育方法の効果に焦点を当てることは，最も重要な問題を覆い隠してしまう。近年影響力を持っている教育効果研究にもこの問題を見てとることができる。

ハマースレーも，ハーグリーブスと同様に教師の専門職性の向上を重視している。しかし，それに対する教育研究の貢献の仕方についての理解が異なっている。ハマースレーに言わせれば，教育研究は，当該の状況においてどの指導方法を採ることが適切かを明示するものとはなり得ない。彼曰く，教育研究の貢献のあり方は啓発的なもの（the enlightenment model）に近い。すなわち，教師が状況を認識する際に依拠する仮説を修正するようなものに近い。

結論においてハマースレーは，現在の教育研究の抱える課題の解決に対し，ハーグリーブスの提案する処方箋は研究を死に至らしめるおそれがあると指摘した。さらに彼は，力のある研究者を失い，教師の専門性も崩壊しかねないとも批判した。

論文の後段で，彼は次のように自らの役割を説明した。すなわち，教育研究は，学校での授業の改善に対してきわめて限定的な役割しか果たし得ない。研究は問題に光を当て，それを分

析し，そしてその問題に対する理解を深めていくことはできるが，学校教育の失敗への対処は，教師の専門性に委ねられるべき実務課題である。それは広く入手可能なエビデンスに還元できるようなものではない。彼はこのようにハーグリーブスの議論を批判した。

4 教育学における議論との共通性

ハマースレーの批判から分かるのは，教育のエビデンスを得ようとする研究（＝議論を簡明にするために「エビデンス研究」と名付ける）を提唱するハーグリーブスとの，教育研究に対する見解の違いである。ハマースレーに言わせれば，エビデンス研究はある特定の変数が教育効果を規定するという前提で考えるが，教育の良し悪しは教師の専門的な判断にゆだねられるものであり，外在的に要因を特定することはできない。ハマースレーは，研究がエビデンスを明示できないことはすでに経験済みであり，だからこそ教育研究は，1960年代後半に質的研究に向かったのではないかと批判した。そして，質的研究は問題の性質を反省的に解明することを目指し，その知見に基づいて実践者の気づかない理解のあり方へと彼らを導くことが使命だと主張した。換言すれば，彼は，研究がなすべきことは，エビデンスの提供ではなく，教師に対して自らの問題理解を反省的に捉えなおさせる知見を提供することだと論じた。

以上からうかがえるのは，過去の遺物であったはずのエビデンス研究の研究枠組がリバイバルしてきたことに対するハマースレーの違和感である。1960年代末にエビデンス研究の問題を指摘して質的研究への転換をリードしたハーグリーブスが，1996年になってエビデンス研究への転換を声高に叫んだことに，ハマースレーの憤りが透けて見える。彼は，ハーグリーブスは質的研究が研究の蓄積を途切れさせたと批判しているが，質的研究が行ったのは理論的発展であると指摘している。エビデンス研究に対する質的研究者の批判のポイントの１つはここにある。つまり，エビデンス研究への転換を求める声は，質的研究者が1960年代末以降に議論を積み重ねてきた理論面での検討の過程，とりわけ表象の危機以降，質的研究者が自己批判しながら研究のあり方を模索してきた過程を無視しているのではないかという点にある。

また，本稿の冒頭で触れたように，ハマースレーは，エビデンス研究を批判する際に，教えるという営みやその担い手である教師に関する議論を参照した。たとえば，彼が引用しているシュワブは，日本でも教育学者の佐藤学らが盛んに紹介しているカリキュラム研究者である（佐藤，1996）。ハマースレーが引用した論文で，シュワブはカリキュラムを編成する際に，既存の理論に依拠する理論モードの弊害を説き，現場に即した実践モードに変換していく必要を説いている（Schwab, 1969）。シュワブによれば，理論モードでは，大規模な集団の成員に適用可能で，長期間の効果をもつとされる知見が示される。しかし，理論モードは，実際の学校現場で起きている生徒の問題を解決するためにはきわめて重要な，誰に対して何をどのように教えればいいのかという課題を考慮することが難しい。これに対して実践モードが提示するのは，教師の行為をガイドするための「判断」である。それは真実とか信頼性が高いといったものではなく，いくつかある選択肢の中から，相対的にどれがより良いかの判断でしかない。

実践モードによるカリキュラム編成は，ルールの当てはめではない。シュワブは，そもそも我々は問題があるとは気づいていても，どのような問題かを理解していないし，また，何をそこで目指すのか，何を必要としているのかも曖昧であり，状況のどこが歪んでいるのかも分からないと述べた。それらは状況の中で徐々に分かってくることであり，それから初めてデータ

を集め，議論していく。シュワブは，どのようなデータが必要なのかは，問題が徐々に形となり浮かび上がる中で初めて分かってくるものだと指摘した。

こうしたシュワブの議論に基づけば，実践と関係なくエビデンスが収集されて，それに基づいて，ある指導方法が上から指示されることは，彼が批判した理論モードの「理論」の補強でしかない。そうしたエビデンスの提示は，理論の信頼性を高めることとなり，実践モードから見れば，よりやっかいな事態である。

また，専門職の役割について論じたショーンにより，1980年代に反省的実践家という概念がもたらされた（Schön, 1983, 訳, 2007）。ショーンによれば，わざや技法に依拠していた医者やエンジニアなどの業務は，19世紀後半から20世紀前半にかけて，科学に基づく技術的実践にとってかわった。さらに第二次世界大戦後に研究費が増大すると，医学モデルと工学モデルが威信を獲得し，そのモデルを教育や社会福祉などの領域においても適用としようとした。しかし，1960年代から80年代になると，直面する課題の複雑性や不確実性，不安定さ，独自性，価値葛藤といった，現実の営みの重要性が気づかれるようになり，技術的合理性の限界があらわになってくる。ショーンが提唱したのは，そのような複雑な状況において，実践家が行為の中での省察により，自身の実践に潜む暗黙の理解に気づくことである。そうした省察の行為が不確実で独自な状況に対して新たな理解の付与を可能にする。ショーンの反省的実践家の概念は，教育界に広く受容され，1980年代以降の各国の教師教育改革に大きな影響を与えた。

このように見ていくと，ハマースレーの質的研究の立場からの批判は，教育学におけるカリキュラムや教師に関する議論に通じるものだったと言える。つまり，ハマースレーの批判は，それだけ教育という営みの特性を踏まえたものだったと言える。

5　エリオットによる事例研究の提唱

次に，エリオット（Elliott, 2001）の議論を紹介しよう。彼は，質的研究の立場からのハマースレーの批判を受け止めつつも，ハーグリーブスが目指した研究の実践への貢献を重視し，両者とは異なる第三のあり方として事例研究の意義を提案した。彼がこの論文を著す前に，ハーグリーブスはエビデンスに関する新たな論文（Hargreaves, 1999）を著した。エリオットは，この1999年の論文においてハーグリーブスは教育研究が教育政策に与える影響を過大評価していると指摘した。その際，エリオットは，マッキンタイア（MacIntyre, 1981, 訳, 1993）の議論を参照し，政治家や官僚が社会工学による専門的知識・技能を強調するのは，科学的に社会を統制しているとする歴史上の「仮装」であると指摘した。マッキンタイア曰く，社会において権力や権威が持たされるのは歴史的な達成であって，ある目的の達成のために社会秩序を操作すべく収集された科学的エビデンスの帰結ではない。エリオットは，教育研究が教育政策に対して重要な貢献をなしうるというハーグリーブスの説明は，政策立案や政策的介入に対する教育研究の効能に関するフィクションに基づいたものだと批判した。エリオットによれば，経験的調査による一般化という手続きで得られた知見が証明しているのは，人々の暮らしには予測不能性が常に伴っているという事実であり，それゆえに政治や官僚制のコントロールに対して，教師には一定の自律性が与えられる必要がある，とのことである。

このような批判的検討に基づいて，エリオットは社会工学的な視点から教育研究の再編成を提案したハーグリーブスとも，オーソドックスに研究の啓蒙的意義を強調したハマースレーとも異なる第三の道を提案した。ここで彼が推奨したのが，真の意味での教育の目標の実現のた

めの，事例研究によるエビデンスの収集である。

このことを説明するために，彼はまず教育哲学者ピーターズの論に依拠して，教育研究が貢献する対象である教育実践の目標について論じた。エリオットはピーターズの言葉を借りて，「教育はそれにとって外因的な価値をもつ事柄に対して手段として役立つ価値中立的な過程であると考えるのはあまりにも安易にすぎる（Peters, 1966, 訳, 1971：28）」と指摘した上で，このような外的な達成目標（ends）は，教育の目標（aims）ではなく，教育の目的（purposes）と呼ぶ方が適切だと主張する。その上で彼は，さらにピーターズの議論に依拠して，教育の第1の目標は，人が教育された状態を意味づける規範や価値に関連しており，そのような状態を可能にする教育的なプロセスを実現することだと指摘した。

エリオットは，このような議論に基づいて，教育の発展のために活用できるとハーグリーブスが想定しているエビデンスは，道具的な効果に関するエビデンスに限定されていると批判した。教育の目標は，その過程や原理に内在するとしたピーターズの指摘を踏まえれば，当該の教授・学習過程が「教育的な」プロセスであるかどうかを証明するためのエビデンスを集めなくてはならない。教育研究の主たる役割は，教室の教育活動と学習効果との統計的関連を明らかにすることではなく，特定の実践の文脈における，一貫性のある「教育的な」営みが可能になる条件を究明することにある。このような議論を経て，彼は，教育の価値や原理の非限定的な性質や，教育方法の文脈依存性ゆえに，「教育的な」研究は事例研究の形式をとるべきだと主張した。その研究は，教育的な実践の「内側」に関して，外部の研究者と教師が協力して遂行されるものである。それは，単なる教育についての研究（research on education）ではなく，「教育的な」研究（educational research）である。そして，エリオットは，研究の知見を現場にあてはめるのではなく，研究の計画や実施の過程にも教師が関与すべきだと説いた。

このようにエリオットは，研究によって示されるエビデンスが貢献すべき教育の目標は何かという点を反省的にとらえなおすことにより，教育研究の独自の貢献のあり方を提示したのである。

6 「エビデンスに基づく教育」の希求にどう対応すべきか

最後に，これら2人の質的研究者の議論を踏まえて，エビデンスに基づく教育が強く求められるようになった日本の現状にどう対応すべきかについて考えてみたい。今の世界的潮流が芽生え始めた20年前の論争からうかがえるのは，「エビデンスに基づく教育」へと舵を切ることは，「教育研究の課題とは何か」，また「教育実践に対して教育研究が貢献できることはどのようなことか」という教育研究の本質に関わる重大事だということである。それは同時に「教育とは何か」，「教師の役割は何か」という教育に関する本質的な問題にも関わっている。

質的研究の立場から見ると，現在の「エビデンスに基づく教育」の隆盛は，1960年代末以降の研究動向を無視し，その後の研究の理論的発展を軽視するものに映る。現在の流れに安易に依拠してエビデンスを得るために研究を発展させていくことは，ハマースレーが言うように，教育研究を死に至らしめるおそれをはらんでいる。これに対してエリオットは，そのリスクを回避するために教育の目標を再確認した上で，教育研究が教育実践に対して貢献し得る可能性を見出し，事例研究の推進を提案した。ただし，そうはいっても，エリオットも教育政策に対して教育研究が役立つとは考えていない。エビデンスや科学的合理性に基づいて推進されるのはフィクションに過ぎないと，彼も考えた。

「エビデンスに基づく教育」を求める声を受け

とめて，教育研究の社会的貢献をアピールするには，エリオットの提案する事例研究は有効な対応策だと見ることができる。それは教師や現場関係者と協働して進められるアクションリサーチの形式をとるものだと想定された。

筆者自身も教育現場の問題に研究者が取り組むための方法論の１つとして，エリオットと同様の参加型アクションリサーチの推進を提唱した（酒井，2014）。その協働のあり方は，具体的に言えば，杉万（2006）が提案した二次モードのアクションリサーチである。それは「質的方法としての解釈的エスノグラフィーを用いながら，一方で社会構成主義的な視点を携えて『気づかざる前提』に気づき，それまでの問題理解のされ方や実践のあり方に対して，新しい可能性を拓いていく（酒井，2014：52）」ものである。

ただし，それはあくまで教育研究の社会的貢献のあり方の1つでしかない。ハマースレーが言うように，研究が実践に役立つ範囲は限定的であり，また，研究そのものの研究関心で遂行される研究も担保される必要がある（酒井，2017）。さらに，教育現場の視点から言えば，教育実践は基本的に教師の専門的な判断にゆだねられており，それゆえに教師の自律性がある程度担保される必要がある。

そのような様々な制限の中でも，教育研究は「エビデンスに基づく教育」という現在の社会の動向に対応することができるし，またその可能性を追求することが必要だろう。なお，再度確認しておきたいことは，エリオットが述べたように，教育研究がエビデンスとして扱うのは，通常言うところのエビデンスとは異なるものである。質的研究の立場にいる研究者は，「エビデンスに基づく教育」の動向に本質的な部分で批判的であり，その動向を支える基本的な教育観や研究観を反省的に捉えなおすことで，独自の貢献をなしうると考えている。

現在の「エビデンスに基づく教育」の動向に対し，質的研究の立場からは，このような緊張に満ちた関わり方での関与を提案することとなる。筆者は，先に紹介した本の中で，このことを「批判を介した規範性（酒井，2014：204）」と命名した。それは，分析的に検討しつつ，その中で常に教育のあり方，教育問題の捉え方，子どもの成長のあり方についてあるべき姿を追究する姿勢である。教育現場に対して教育研究が果たしうる役割はそのようなものであり，教育政策を先導したり，実践に対して有効な方法論を指し示すようなものとはなり得ない。教育の政策や実践にきわめて禁欲的に関わることが，教育研究者に与えられた役割だと思うのである。

注

1) https://professorjohnelliott.wordpress.com/を参照。

文献

Elliott, J.,2001 “Making Evidence-based Practice Educational”, *British Educational Research Journal*, 27(5):555-574.

Hammersley, M., 1989, Classroom Ethnography. Milton Keynes, UK: Open University Press.

————, 1997, ”Educational Research and Teaching: A Response to David Hargreaves' TTA Lecture”, *British Educational Research Journal*, 23(2):141-161.

Hammersley, M. & Atkinson, P., 1983,

Ethnography: Principles in Practice, London: Tavistock.

Hargreaves, D.H., 1967, *Social Relations in Secondary School*. London: Routledge and Kegan Paul.

————, 1996, "Teaching as a research-based profession: possibilities and prospects", *Teacher Training Agency Annual Lecture 1996*, London: Teacher Training Agency.

————,1999, Revitalising educational research: Lessons from the past and proposals for the future. Cambridge Journal of Education, 29(2): 239–249.

岩崎久美子,2011,「教育研究エビデンスの課題－知識社会における産出・普及・活用」『国立教育政策研究所紀要』第140集:95-112。

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 —— 教育の標準化・市場化の中で」『教育学研究』82 (2): 216-228.

MacIntyre. A.C., 1981, After Virtue: a study in moral theory, London: Duckworth.（篠崎栄訳,1993,『美徳なき時代』みすず書房.）

Oancea, A., and Pring, R., 2009, "The importance of being thorough: On systematic accumulations of "what works" in educational research," in D. Bridges, P. Smeyers, and R. Smith (Eds.), *Evidence-based education policy: What evidence? What basis? Whose policy?* : 11-35, .Malden, MA: Wiley-Blackwell.（籾井圭子訳, 2013,「徹底的であることの重要性」柘植雅義ほか訳『エビデンスに基づく教育政策』勁草書房:17-48.）

Peters, R. S., 1966, Ethics and Education, London: Allen & Unwin.（塚崎智, 三好信浩訳『現代教育の倫理 —— その基礎的分析』黎明書房, 1971.）

酒井朗, 2014,『教育臨床社会学の可能性』勁草書房。

————, 2017,「教育社会学と教育現場」日本教育社会学会編, 本田由紀, 中村高康責任編集『教育社会学のフロンティア 1 学問としての展開と課題』岩波書店。

佐藤学,1996,「ディシプリン中心カリキュラムの継承—シュワブの「構造」概念を中心に」佐藤学『カリキュラムの批評　公共性の再構築へ』世織書房：83-94。

Schwab, J. J., 1969, "The practical: a language for curriculum", *School Review*, 78：1-24.

Schön, D. A.,1983, *The Reflective Practitioner: How professionals think in action*. London: Temple Smith.（柳沢昌一・三輪健二監訳,2007,『省察的実践者とは何か —プロフェッショナルの行為と思考』鳳書房.）

杉万俊夫, 2006,「質的方法の先鋭化とアクションリサーチ」『心理学評論』49 (3)：551-561。

玉井航太, 藤田英典,2017「エビデンスに基づく教育のための縦断データの解析方法」『教育研究』国際基督教大学,59：5-16。

Winch, P., 1958, *The Idea of a Social Science and its Relation to Philosophy*, London; Routledge（森川真規雄訳, 1977,『社会科学の理念 - ウィトゲンシュタイン哲学と社会研究』新曜社.）

特集論文

6

エビデンス・ベースド・アカウンタビリティの批判的再構築はいかにして可能か

澤田 稔
上智大学総合人間科学部 教授

1 はじめに 本考察の見通し

本稿は本誌に掲載されているものの，社会学的というよりも教育学的な考察である。本考察は，基本的には，カリキュラム・教育方法論に依拠し，エビデンス・ベースという要因を，各学校現場に求められるアカウンタビリティという要因に結びつけて一定の知見を明らかにしようとするものである。

他方で，本稿が社会学との接点をまったく持たないわけではない。本稿は，教育社会学でも参照される批判的教育学（ネオ・マルクス主義思想の系譜につらなり，かつて「新しい教育社会学」と呼ばれた研究業績群とも関連の深い北米の教育学的研究）の視座に根ざした考察だからである。批判的教育学における近年の議論では，アカウンタビリティやエビデンス・ベースドという観点に関して，それらが新自由主義的な，あるいは親市場的マネジメント論が公的領域に侵入したものとして否定的に言及されつつも，その問題点を克服できるような代替案の可能性が模索されてきた。

そこで，本稿もこの延長線上で，新自由主義のある側面や市場原理主義に対する批判的視角を維持しつつも，エビデンス・ベースド・アカウンタビリティを全面否定するというスタンスは採用せず，その不可避性を認めた上で，一定の公正性を担保できる方法の探求を目指す。よって，その問いを端的に表現するなら「より公正な学校教育に向けたエビデンス・ベースド・アカウンタビリティはいかにして可能か」となろう。

本稿では，この問いへの解に接近すべく，ある一つの学校の具体的な事例を，エビデンスに基づいてアカウンタビリティを果たす方法の一つのモデルとして参照することで議論を進めたい。その事例とは，ボストンの公立学校制度の中に特別に設けられた「パイロットスクール（pilot schools）」と呼ばれる枠組とここに属する学校に求められるスクール・クオリティ・レヴュー（School Quality Review：以下SQRと表記）と呼ばれる特別の評価制度，およびパイロットスクールの一つであるミッションヒル校（Mission Hill School：以下MHSと表記）の取組である。

以下では，最初に，批判的教育学におけるアカウンタビリティ論の骨格を，主としてアップル（Apple, M. W.）の議論に依拠して整理し，そこで肯定的な見通しをもって示唆されているアカウンタビリティ論の可能性に光をあてる。次に，その新たなアカウンタビリティとその果たし方の具体的なあり方について，上述のパイロットスクールという制度的枠組とSQRの仕組の概要を確認した上で，MHSの特質および同校の最近のSQRへの対応を簡略に振り返り，より公正な学校教育の実現という視点から，SQRで求められるようなエビデンス・ベースド・ア

カウンタビリティが持つ意義と課題をまとめたい。ただし，紙幅の都合上，SQRにおける報告書の内容詳細を考察対象に含めることはできず，ここでの議論に必要な限りの簡単な紹介にとどまるだろう。筆者は2011年3月に同校を初訪問以来，毎年ほぼ２回各１週間ずつの訪問調査を重ねて来たので，ここでの議論はその機会に得た知見を背景としているが，本稿ではそこで得た観察・聞き取りデータを参照する余裕がほとんどないことを予めお断りしておきたい。

2 批判的教育学における アカウンタビリティ論

批判的教育学におけるアカウンタビリティ論の中軸は，標準化されたテストに基づいて学校に求められる新自由主義的なアカウンタビリティの持つ問題点の剔抉（てっけつ）にある。そこでは，教育の市場主義的な分権化モデルと，カリキュラム統制・学力試験・アカウンタビリティの三位一体的な集権化モデルとが矛盾なく結びつくことが指摘される。市場化と「選択」の自由の行使には，「消費者」が商品・サービスを比較・選択するために必要な物差しとして標準化されたデータが重用されることになるというわけである（Apple, 2006 : 91）。

ここで特に批判の対象となっている学力試験は，ハイ・ステイクス・テスト（high-stakes testing）と呼ばれる。NCLB法（どの子も置き去りにしない法）はもはや過去のものとなったが，それでも各州の統一学力試験が，全公立学校に実施が義務付けられているだけでなく，学校別の結果やランキングが公になり，公式の学校評価の最主要項目となる場合が多いだけに，依然として，各学校にとって賭金が高く，大きな圧力を与えるものであり続けている。

不平等の再生産に関する社会学的知見に依拠する批判的教育学は，こうした学力試験の結果が，学校が抱える家庭・地域の経済的・文化的諸条件に大きく左右されるものであるとして，これに教育評価上過度の比重が置かれることに対しては極めて否定的である。アップル（Apple, M.W.）によれば，「テストの点数を向上させるという目標に最大の注意を払うことは，様々なリソースや教職員の経験，税基盤の有無や税金面でのサポートの差，貧窮状態の有無，リスペクトや生活資金を得られる仕事の欠如，貧しい住居や不十分な医療等々に関して，貧しい地域と豊かな地域との間に歴然と存在する不平等から目をそらすことにな」るからであり，「こうした問題から注意が逸れることによって，教育上の不平等の原因となっている問題を解決することが一層困難にな」るからである（2006: 96）。

アップルは，イギリス出身の政治経済学者レイズ（Leys, C.）の議論を参照しつつ，こうしたアカウンタビリティの様態が，監査文化（audit culture）と呼ばれるものと密接に結びついている点も指摘している。すなわち，公的機関・公務員の業績測定・評価において，ビジネス由来の独立した監理（independent supervision）に関する諸概念の使用が激増し，数値による測定・評価の圧力が広範囲に高まるのだという。公的部門に大きな信頼を寄せ，集合的資源が持つ効果やその適切な利用を保証する民主的な力を有する市民からなる社会に替わって，監査や査定の準備に汲々とする「被監査者」からなる社会が登場したのだと。そこに見られるのは，各大学・学校・病院等の相対的効率性を示すと称して用いられるランキング表（league tables）の跳梁跋扈（ちょうりょうばっこ）であり，これに基づく懲罰的文化である。公的査察機関は個々の「失敗している」教員や学校などを「名指し辱める」という仕事に加担し，その「失敗している」諸機関に取って替わるように招き入れられるのが民間企業だという図式である（Apple, 2005）。

このような背景理解を示す批判的教育学者と

してのアップルは，学校に対するアカウンタビリティの要求に際して学力テストの結果に置かれている比重が過剰であるとの現状認識に基づき，それを打開する可能性を持つ別方向のアカウンタビリティ論を模索している。彼は，その契機としてこの言葉の語源に着眼した知見に注目する。それは，アカウンタビリティの語源が（古フランス語やラテン語にまで遡ると）「計算する」だけでなく「物語る」という意味とも関係しているにもかかわらず，この言葉をもっぱら前者の意味に還元することで，後者の語源にダメージが与えられているという知見である（Apple, 2006 : 92）。

こうした知見から導かれるのは，数値に還元されないより質的なデータに基づく対抗ヘゲモニー的アカウンタビリティ論である。その例の一つとしてアップルが取り上げているのは，政治的にリベラルな理念を掲げ，標準化されたテストに対する批判的代替案としての学習評価のあり方を提言している進歩主義的教育団体フェアテスト（Fair Test=The National Center for Fair & Open Testing）による評価論である。このアプローチにおいては「標準化されたテストが読み書きと算数に関して用いられているものの，その使用は限定され，生徒による様々な発表や作品，展示，ポートフォリオ，パフォーマンスなどの情報によって埋め合わされている。そこでは個別的で断片的な知識にとどまらない，広く定義された資質・能力に力点が置かれ，コミュニティの価値や機会のあり方の内部でオーセンティックな状況に埋め込まれた比較的緩やかでローカルなカリキュラム要件に依拠することにより，保護者やコミュニティの側のより民主的な関わりを可能にするだけでなく，より確固たるオーセンティックな評価形態の使用による改革の余地が与えられる」のだという。こうした評価論は，米国では，1990年代後半頃から注目されるようになったものである（Apple, 2006: 118；澤田，1997）。

アップルは，この他にも自ら編著に関わった実践論文集『デモクラティック・スクール』に触れ，「公共的目的や実践に関する様々な説明の仕方が民主的公共圏において熟議されることを可能にするような公共的アカウンタビリティ」の可能性について言及しているが，彼の議論はこれら以上に具体的に詳細なアカウンタビリティのあり方に踏み込んではいない（Apple, 2006; 120-2）。では，こうした批判的教育学が目指した対抗ヘゲモニー的アカウンタビリティの具体的実践可能性はどこに見出すことができるだろうか。

3　ボストンのパイロットスクールとSQR

3-1　パイロットスクールという制度的枠組

冒頭で触れたように，MHSはボストン学区内の公立学校に属するが，従来型の一般的な公立校とは異なり，パイロットスクールというカテゴリーに含まれる。パイロットスクールとは，チャータースクールと従来型の一般的な公立学校との中間的な存在である。チャータースクールは全米に数多存在するが，パイロットスクールという制度は，ボストン，あるいはマサチューセッツ州に固有の実験的な試みとして始まった。

チャータースクールは，1991年に全米で初めてミネソタ州で，マサチューセッツ州では93年に法制化されたが，公設民営型学校とも呼ばれ，州による認可を受けた上で，一般の公立学校と同等の運営資金が生徒数に応じて州政府から公的に投入されながらも，州や市の様々な規制に縛られない運営が可能な学校で，民間営利企業も少なからず経営に乗り出している。比較的容易に設立でき，カリキュラム・授業日数や人事・勤務時間等に関して自由な管理運営が許さ

れており，当該教育行政区外在住の児童・生徒の在籍も可能だが，アカウンタビリティが求められ明確な結果責任が問われる。

これに対して，パイロットスクールは，部分的にチャータースクールと類似の特徴を有するものの，あくまで正規の公立校に属し，予算措置は他のすべてのボストンの公立校と同じく市の財源から同様の予算配分・措置がなされ，在籍児童・生徒の家庭はボストン学区内に在住していなければならない。また，チャータースクールの教員が教職員組合に所属していることは稀だが，パイロットスクールの教員は他の公立校同様ほぼ全員が組合員で，給与面でも組合が交渉の末に獲得した給与水準が保証されている。他方，勤務時間など一部の労働条件に関しては，組合が順守を求める水準に必ずしも縛られない。

パイロットスクールが，他の一般的な公立校と大きく異なるのは，チャータースクールと類似の以下の５つの点に関する自律性が与えられていることである。すなわち，①予算使途・配分（総額のみが決められている），②教職員雇用（ボストンに登録されている正規教員以外からも雇用が可能），③カリキュラム（州統一学力試験の実施は義務だが市のカリキュラム基準に従う必要がない），④ガバナンス（校長の任免権限さえ有する運営評議会を置くことも可能），⑤授業日数・時数・日程等に関する裁量権，である（BPS , 2009）。

このようにチャータースクールと一般的な公立学校との中間的なカテゴリーが設置された背景には，前者が法制化された頃，後者において（人口動態的理由と公立以外の学校を選択する保護者が増えたという理由により）就学率の低下傾向が見られたという事情があった。ボストン教職員組合は，1993年，公立の学区内チャータースクール（in-district charter schools）とも呼ぶべきパイロットスクールの設置を含む市当局との契約交渉に応じ，翌94年には合意，95年には最初の６校が開校した。現在，ボストン全公立校125校中21校がパイロットスクールで，その就学者数は全体の１割強となっている（Perlman, 2000 ; Clinchy, 2000）。

3-2　エビデンス・ベースド・アカウンタビリティとしてのSQR

パイロットスクールは，運営上の自由裁量範囲が一般の公立学校に比べて大きい一方で，学校独自の取組が有意味な効果をもたらしているかどうかについて相応のアカウンタビリティを果たすことが求められる。このことはチャータースクールにも該当するが，チャータースクールの場合には，原則として州の統一学力試験の結果が最も主たるエビデンスとして扱われるのに対して，パイロットスクールにはこれとは異なるエビデンスに基づくアカウンタビリティの果たし方が認められている。その仕組がSQRである。これはスクール・セルフ・スタディと呼ばれる自己点検・自己評価報告書（以下，SSS）と，SQRチームによる３日間連続の学校訪問調査が実地視察及び各代表者への聞き取りに基づいて行われ，各パイロットスクールとも5年毎に実施される（BPS, 2014）。

こうしたアカウンタビリティ・評価の仕組を確立する上でも，各パイロットスクールがSQRを進める過程でも，重要な役割を果たしてきたのが協働教育センター（Center for Collaborative Education：以下CCE）である。CCEは，パイロットスクール網の親組織とも呼ばれ，その運営やカリキュラムづくり・学習評価等に関して，様々な助言・研修機会を提供する機関として，1995年に設立された非営利組織である。CCEは，パイロットスクールに専門的技術支援を提供するという契約をボストンの教育局と結んでいる。他方で，CCEは，MHSもその加盟校であるリベラルで進歩主義的な教育団体として全米的に有名なエッセンシャルスクール連盟（the Coalition of Essential Schools）のボストン支部的な性格を帯びており，いわゆるオーセンティッ

ク評価の導入・推進に尽力してきたという経緯をもつ（Nathan & Myatt, 1998 ; CCE, 2004）。

このCCEが，市の委託を受けて，SSSのための書式やルーブリックを作成し，その報告書や訪問実地視察に向けた説明会を実施している。その手順は ①CCEによる学校教職員全体への説明会でSQRの目的や手順・スケジュールを確認，②これを受け学校による4〜6人のSSS作成代表者チームの編成，与えられたルーブリックに基づく自己点検自己評価の開始，③ 学校側の希望やCCEの意見を反映した4〜6名のSQRチーム（教育を専門とする外部評価者チーム）の編成，④学校側によるSSSの完成，⑤CCEによるSQRチームへの説明会の実施，⑥SQRチームがSSSに目を通した上で3日間連続の学校訪問による実地視察・聞き取り調査の実施，⑦2〜4週間でSQRチームによるSQR報告書の完成，⑧学校側からSQR報告書への回答書提出，⑨市の教育長及び教職員組合執行部によるSQR報告書と学校側回答書の閲覧・検討，⑩学校側によるSQRに基づく改善案の開始（BPS, 2014）という経過を辿る。

SSSは①ヴィジョンとミッション，②カリキュラム・教育方法・評価，③専門性開発（PD）・向上，④保護者・地域住民参加，⑤生徒支援，⑥リーダーシップ及びガバナンスという６分野に関して作成され，各分野にはさらに各々10件前後の細目点検内容が評価の観点として設けられており，それぞれの評価項目・観点に対して4段階評定のためのルーブリックが与えられていると同時に，各分野・細目に関して記述式の詳細な説明，および，それらの自己評価の妥当性を示すための資料をエビデンスとして添付することが求められる。

SQRの核となる訪問・実地調査は，３日間のみとはいえ連日午前8時から午後5時までわずかな休憩を挟みつつ，授業および各種会議・研修の視察や，各立場・学年の教員から子ども・保護者・地域住民の代表者に対してまで聞き取りが実施され，これらに基づく評価報告書が執筆される。その結論部分では，学校の長所と課題，及び改善に向けた勧告が詳細に示され，学校側は，その内容を踏まえて長所の維持と課題の克服のために具体的な施策を講じていくことが求められるのである。

4 まとめ
MHSによるSQRへの対応から見た エビデンス・ベースド・アカウンタビリティの可能性

MHSは1997年に，進歩主義的教育家として名高いマイヤー（Meyer, D.）により設立された幼小中一貫校で，現在3〜14歳の子どもたち約250名が在学している。ほぼ全て複式学級で，各学年2〜3クラスで構成され，１クラスあたりの子どもの数は20名以下，各クラスには担任以外の補助教員が配属されている。入学は抽選制だが，多様な社会階層の子どもたちがおり，特に必ず一定数の社会経済的に恵まれない家庭出身の子どもが含まれるように在学生が構成されている。また，在籍する子どもたちの半数近くは何らかの特別な支援を必要とし，様々な障がいを持った子どもが持たない子どもと一緒に学ぶフル・インクルーシブ教育が実践（ただし，特別支援のための通級教室が常設）されている。MHSでは，このように様々な点で社会的に不利な条件を抱える子どもを含めて一人ひとりが包摂される学習空間の構築が目指されている。

MHSが多く持つ注目すべき特質の一つを，その中心理念とカリキュラム・教育方法にフォーカスして述べるなら，次のようになろう。すなわち，「民主主義という言葉に潜む理想を堅持し，全ての人々が平等な存在として尊重されるべきだと訴える」ことができる次世代の若者を育てるという理念に基づき，民主主義社会を支える市民に必要な批判的思考の一種としての「精神の習慣（Habits of Mind）」（表1）を少しでも内面化・意識化できるよう，一人ひとりの子どもが意欲的に学習に取り組むことがで

表１　精神の習慣（Habits of Mind）

証拠・証明（Evidence）	自分が知っていることは，どうすればわかるのか。どの程度の確証があるのか？
観点(Viewpoint）	これは，誰の視点からの提示されているものなのか。別の観点から異なる理解の仕方はできないだろうか？
連関(Connections）	これは，私たちがすでに知っていることとどうつながっているのだろうか？一定のパターンはあるだろうか？
妥当性(Relevance）	このことはなぜ重要なのだろうか。どのような意義があるのか？
推測(Conjecture）	このことは，もし事情が違っていればどうなるか。異なる過去・現在・未来を想像できるだろうか。

き，この思考形式を様々な場面で活用するプロジェクトベースのテーマ中心カリキュラムに学校全体で取り組んでいる点にあると（Meier & Knoester , 2017: 124）。

他方で，MHSは，州統一学力試験の成績では目立った結果を全く残せていない。にもかかわらず，同校が持続可能な発展を遂げていることは否定しようがない事実である。それは，設立以降，校長の交代を経ても，安定した入学希望者数，学校運営が継続していることに表れている。このことは州の統一学力試験の結果が依然ハイ・ステイクスな状態にある米国の現状に照らすと容易に当然視できない。

こうした事情に鑑みると，MHSの持続可能な発展性を支えている要因としてパイロットスクールという制度的枠組と，SQRという質的なエビデンスに基づくアカウンタビリティの仕組が公的に確立されているという諸点を無視することはできない。MHSは，パイロットスクールとして，学校運営上チャータースクール並みの裁量範囲が与えられているからこそ，上述のような特徴的な理念と運営方法を維持することができているのであり，学力試験とは別の質的なエビデンスによって学校の存在理由・価値を公式に検証・審査するシステムが確立されてきたからこそ，学力試験では測定できないが，社会的公正性という点に照らして一定の積極的評価を与えられてしかるべき上述のような特質を持つMHSの存在が維持されてきたと考えることができるからである。

この点は，マイヤーがMHSよりも以前にニューヨークに設立したセントラル・パーク・イースト中等学校（Central Park East Secondary School）がもはや名前しかのこっていないこと（Apple & Beane, 2007），同じくセントラル・パーク・イースト第１小学校（Central Park East 1 Elementary School）が最近経験した危機的状況とは対照的である（Zimmer, 2016）。ニューヨークにはパイロットスクールという制度はなく，SQRのような評価システムも導入されていない。

ただし，パイロットスクールやSQRのような仕組が持つ両義性に関しては注意が必要だろう。

それは，第１に，ルグラン（Le Grand, J.）が定式化した「準市場（quasi-market）」というファクターと関係しているように思われる（2009）。SQRというエビデンス・ベースド・アカウンタビリティの様態は，批判的教育学でもその可能性が模索されたのと同じ方向の質的でオーセンティックな評価方法に通底するものであったが，それはパイロットスクールという，批判的教育学においてはむしろ否定的に論じられてきた「選択と競争」の論理に基づく公共サービスの制度設計と相即不離な仕組として導入されたものである。ここには準市場が，より社会的に公正なエビデンス・ベースド・アカウンタビリティを導入・推進する可能性を孕んでいることが示されているが，そこで選択と競争の焦点となるサービスの質が，MHSの重視する類の教育の質を駆逐しない保障はないという

点も確認されるべきであろう。

第２に，SQRが学力試験の結果だけでは評価できない教育の質的側面に照射するという点では新たな教育評価の可能性を拓くものではあるが，他方でアカウンタビリティのあり方の一種である以上「監査文化」という要因と無縁ではなかろう。監査の機会は，学校が自らの教育理念を実現すべくその存在意義と課題を明確化し，次の改善へと結び付けるための手段に過ぎないはずだが，この目的・手段の関係が反転する危険性を想像することは難しくないからである。くわえて，質的なエビデンスがより重視されるとしても，第１の点で見たように，どのような質に価値を置くかは多様であり得るという点も注意すべきだろう。

実際，MHSで教員を務めたのち教育学者となったノースター（Knoester, M.）の考察を参照すると，パイロットスクールの中で，民主主義の理念や平等主義的なインクルーシブ教育を重視する姿勢を明確に表している学校はMHSの他にはほとんど見出せない（2011）。

では，上記のような両義性を孕む制度において，MHSはどう対応してきているのか。MHSは最近ではSQRを2013年度に実施した。その際の学校側の責任者であった現校長ガビンズ（Gavins, A.）は，筆者による聞き取り（2018年3月7日）において「この機会はハイステイクスなので緊張はします。しかし，それはポジティブな意味です。というのも，非常によいフィードバックの機会になるからです。」と述べたように，SQRを学校全体で共有している目標・理念をよりよく実現するための重要な手段として捉え，積極的に対応している。実際，同校長のSQRチームへのオープンな姿勢，及び学校全体で教育理念・目標が分厚く共有されているという事実は，同チームによる評価報告書に明記され高い評価を受けた（BPS, 2014）。

以上を踏まえて，MHSの取組を再解釈してまとめるならば次のようになろう。同校は，自校の理念や教育方針がパイロットスクールの中でも極端な少数派であることを自覚しつつも，準市場的諸条件を積極的に活用して，民主主義の理念に根ざした教育の持続可能な発展を目指していると捉えることができる。すなわち，準市場が「もう一つの見えざる手」によって質の高いサービスの効率的提供を相対的に首尾よく達成する可能性が高いシステムではあるとしても，MHSはそのシステムに依存するのではなく，自らの理念を実現するためのアクティヴィズムとして，質的なエビデンスに基づくアカウンタビリティを積極的に果たそうとしていると考えられるのである。それは，同校の実践があらゆる面で常に成功していることを意味しない。マイヤーとノースターが，民主主義それ自体を常に未完の作品（work in progress）として捉え，にもかかわらずその理想を維持し，それを具現化していく営みの重要性を強調しているのと同様である（Meier & Knoester, 2017）。こうした困難な実践の問題点や不十分さを批評的に指摘することは容易い。しかし，むしろ今後学問的に必要なのは，そうした困難さを再帰的に引き受けた上で新たな地平を切り開こうとしている一部の教育実践が持つアクティヴィズム的側面を正面に見据えた新たな分析視角を手に入れることではないだろうか。

＊

本稿は，文部科学省科学研究費基盤研究（C）（課題番号16K04624）による研究成果の一部である。

文献

Apple, M. W., 2005, “Education, markets, and an audit culture” *Critical Quarterly*,47(1 - 2):11-29.

————, 2006, *Educating the “right” way: Markets, standards, god, and inequality* (2nd ed.), New York: Routledge.MHS (Original work published 2001).（大田直子訳, 2008,『右派の/正しい教育 市場, 水準, 神, そして不平等』世織書房.）

Apple, M. W., and Beane, J. A. Eds., 2007, *Democratic schools* (2nd ed.), Portsmouth, NH: Heinemann.（澤田稔訳, 2013,『デモクラティック・スクール：力のある学校教育とは何か』上智大学出版.）

Boston Public Schools, 2009, *Pilot schools*. (Retrieved May 1, 2018, from https://www.bostonpublicschools.org/Page/4682)

————, 2014, *2013-2014 School Quality Review Team Report. The Mission Hill School*, Boston, MA. submitted: July 15, 2014. unpublished paper.

————, 2014, *School Quality Review Guide*, Center for Collaborative Education.

Center for Collaborative Education, 2004, *How Pilot Schools Authentically Assess Student Mastery*. (Retrieved May 1, 2018, from http://cce.org/files/AuthenticAssessment_2004.pdf)

Clinchy, E. Ed., 2000, *Creating new schools: How small schools are changing American education*. New York: Teachers College Press.

Knoester, M. , 2011, “Is the outcry for more pilot schools warranted? Democracy, collective bargaining, deregulation, and the politics of school reform in Boston”, *Educational Policy*, 25(3): 387-423.

Le Grand, J., 2009, *The other invisible hand: Delivering public services through choice and competition*, Princeton : Princeton University Press.（後房雄訳, 2010,『準市場 もう一つの見えざる手 ― 選択と競争による公共サービス』法律文化社.）

Meier, D., and Knoester, M., 2017, *Beyond testing: Seven assessments of students and schools more effective than standardized tests*, New York: Teachers College Press.

Nathan, L., and Myatt, L., 1998, “A journey toward autonomy” *Phi Delta Kappan*, 80(4):278.

Pearlman, B., 2000, “Smarter charters? Creating Boston’s pilot schools” Clinchy (Ed.), *Creating new schools: How small schools are changing American education*: 38-48, New York: Teachers College Press.

澤田稔, 1997,「アメリカ合衆国における教育方法改革の最前線」, 松浦善満・西川信廣 編,『教育のパラダイム転換 ― 教育の制度と理念を問い直す』福村出版。

Zimmer, A.,2016, “New Principal Ruining Legendary Progressive Public School, Parents Say” *DNA Info*. (4/13/2016, https://www.dnainfo.com/new-york/20171102/tribeca/students-with-disabilities-support-services-nyc-schools/)