

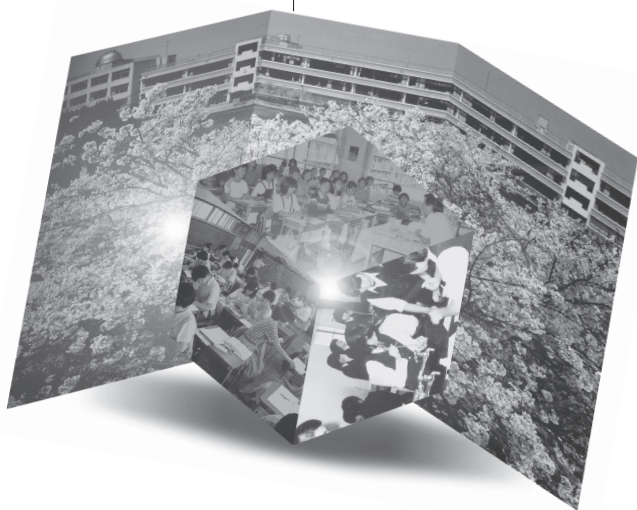
特集

# 調査のフィールド としての学校

文部科学省の「全国学力・学習状況調査」がさまざまな反響を呼んでいる。教育と階層，学校間格差，義務教育の地方分権化，学校評価，「モンスター・ペアレント」等々，学校現場への社会的関心も高まりをみせている。

こうした関心に応えるためには，学校現場についての実態把握や分析が不可欠であり，学校をフィールドとした調査への期待も高まっている。学校をフィールドとした調査はどのような現状にあり，その課題と可能性はどこにあるのか。

特集では，関心と期待が高まる「調査のフィールドとしての学校」について，社会学，教育社会学，テストや測定理論など複眼的な視点から議論します。





## 「学力問題」に切り込む



### 1 はじめに

本稿では、「学力問題」を切り口に、社会調査の可能性と課題について論じてみたい。

筆者はほぼ四半世紀にわたって、学校社会学の領域で仕事をしてきた。その過程の中で、学校を対象とするかなり多くの調査研究に従事してきた。それは、大規模サンプリング調査から長期にわたるフィールドワークにいたるまで多岐にわたる。そして、最近の7~8年は、学力の問題を主たる研究テーマとして活動を続けている。本稿では、そうした筆者自身の経験をベースに、社会調査という方法がもつ可能性とそれに付随する諸課題について、若干の考察を行おうと思う。

以下、次のような手順で、この問題にアプローチしたい。まず、2節では「学力問題」とは何か、を簡単にあとづけたい。そのあとの3節と4節が、本稿の中心となる部分である。3節では、「研究としての社会調査」と題して、欧米の「効果のある学校」論を日本の現実に適用しようとしてきた筆者自身の研究を振り返り、社会調査がもつ可能性について一研究者としての視点からの考察を行う。続く4節では、「施策としての社会調査」と題して、昨年からスタートした「全国学力・学習状況調査」を取り上げ、国（文部科学省）

志水 宏吉（大阪大学大学院人間科学研究科教授）

が主体となって実施する大規模な社会調査が有する課題について検討する。そして、最後の5節ではまとめの考察を行う。

### 2 学力問題

今日の教育界において、おそらく最大の課題として認識されているのが「学力問題」である。いわゆる「学力低下論争」が勃発したのが、1999年から2000年にかけてのことであった（市川、2002、山内・原編、2006）。大学生の学力低下についての「告発」に端を発したこの論争は、「学びからの逃走」論や「ゆとり教育亡国」論といったバリエーションを奏でながら、数年後にピークに達した。筆者らが、「二こぶラクダ化」という形で、子どもたちの「学力格差」の問題に警鐘を鳴らしたのは、まさにそのさ中のことである（荻谷・志水・清水・諸田、2002）。

ほどなく学力低下論争は終結する。筆者らのものもふくめ、いくつかの学力調査データが明らかになるにつれ、やはり子どもたちの学力は低下していると言わざるをえないという論調が支配的になってきたからである。その趨勢を決定的にしたのが、2004年の年末に発表された2003（第2回）PISA調査の結果であった。PISAとは、Programme for International Student Assessment（生徒の学

習到達度調査)の略であり、これは、IEA(国際教育到達度評価学会)が実施するTIMMS(2003年国際数学・理科教育動向調査)などと並ぶ代表的な国際学力テストである。日本の結果は、2000年に実施された第1回の結果と比べてかなり悪化していた。国際的な視点からみて日本の子どもたちの学力には多くの課題があるということが明らかになったのである。

時を同じくするように、文部科学省は、「ゆとり教育」路線から「確かな学力」路線へと舵を切った。今日へとつながる学力重視路線がスタートしたのである。子どもたちの学力をいかに向上させるか(学力水準の問題)、「しんどい」層の学力をいかに下支えするか(学力格差の問題)、そして知識の習得だけでなく活用する力をいかに向上させるか(学力構造の問題)が、現在の教育現場の最大の争点となっているのである。

### 3 研究としての社会調査 ——「効果のある学校」論の適用

筆者が学力研究に着手したのは2001年、まだ学力低下論争が華やかなりし頃のことである。当時の同僚であった荻谷剛彦氏らと自前の学力調査を実施した。その結果については、荻谷・志水編(2004)に整理してある。この調査は、子どもたちの学力の「二こぶラクダ化」を指摘したものとして知られているが、筆者自身が見出したのは、学力の低下や学力格差の拡大の克服に成功している「がんばっている学校」の存在であった(志水, 2003, 2005)。以来筆者は、欧米で展開されてきた「効果のある学校」論の枠組みを日本の学校に適用する仕事を続けている。「効果のある学校」論にはいくつかのバージョンがあるが、筆者らのグループが採用してきたのは、

アメリカのR.エドモンズらが提唱した「教育的に不利な環境のもとにある子どもたちの学力を下支えしている学校」というものである(鍋島, 2003)。最近の仕事として、2006~07年にかけて行った調査研究があるので、それを簡単に振り返っておこう。

データは、2006年度に大阪府教育委員会が府内で実施した学力実態調査である。分析を委託された私たちのグループは、上述の枠組みにしたがった量的分析を行い、「効果のある学校」を同定した。具体的には、「塾に行っていない子どもたち」や「家庭の教育環境にあまり恵まれていない子どもたち」の基礎学力の水準をかなりの程度下支えすることに成功している学校を「効果のある学校」と判定したのであった。その数は、対象校全体の2割程度にのぼった。さらに私たちは、翌2007年度、府教委のサポートのもとに、それらの中から学校規模や地域バランスを考慮したうえで10校(小学校5校・中学校5校)を選定し、「効果を生み出している要因」を探究するための質的調査を実施した。すなわち、各校に対して研究者1名と大学院生1名計2名からなるチームを割り当て、1校あたりのべ25~30回にわたるフィールドワーク(訪問調査)を行ったのである。最終的に、各校で見出された知見をもちより作成したものが、図1に示した「スクールバスモデル」である。

この図は、「効果のある学校」という概念を実践的な意味合いにおいて拡張した「力のある学校」(=すべての子どもをエンパワーする学校)をつくり出すために必要な8つの要因を、「スクールバス」のメタファーを使って立体的・構造的に示そうとしたものである。18名にのぼる研究チームの2年間にわたる協働作業の成果として生み出されたのが、こ

教職員集団の強力なエンジンと学校運営のハンドルさばきをスクールバスの中心とし、生徒指導と学習指導はバスを導いていく前輪、校種間連携と家庭連携は運転を安定したものにす後輪、学校環境・学校文化は、バスのインテリアとボディと考えた。

このスクールバスが走っていく道は決して平坦ではないだろうが、8つの要素をうまく連携させて、少々悪路であっても力強く乗り越えていく学校の姿を思い描いた。

Together	
① 気持ちのそろった教職員集団	(Teachers)
② 戦略的で柔軟な学校運営	(Organization)
③ 豊かなつながりを生み出す生徒指導	(Guidance)
④ すべての子どもの学びを支える学習指導	(Effective teaching)
⑤ ともに育つ地域・校種間連携	(Ties)
⑥ 双方向的な家庭とのかわり	(Home-school link)
⑦ 安心して学べる学校環境	(Environment)
⑧ 前向きで活動的な学校文化	(Rich school culture)

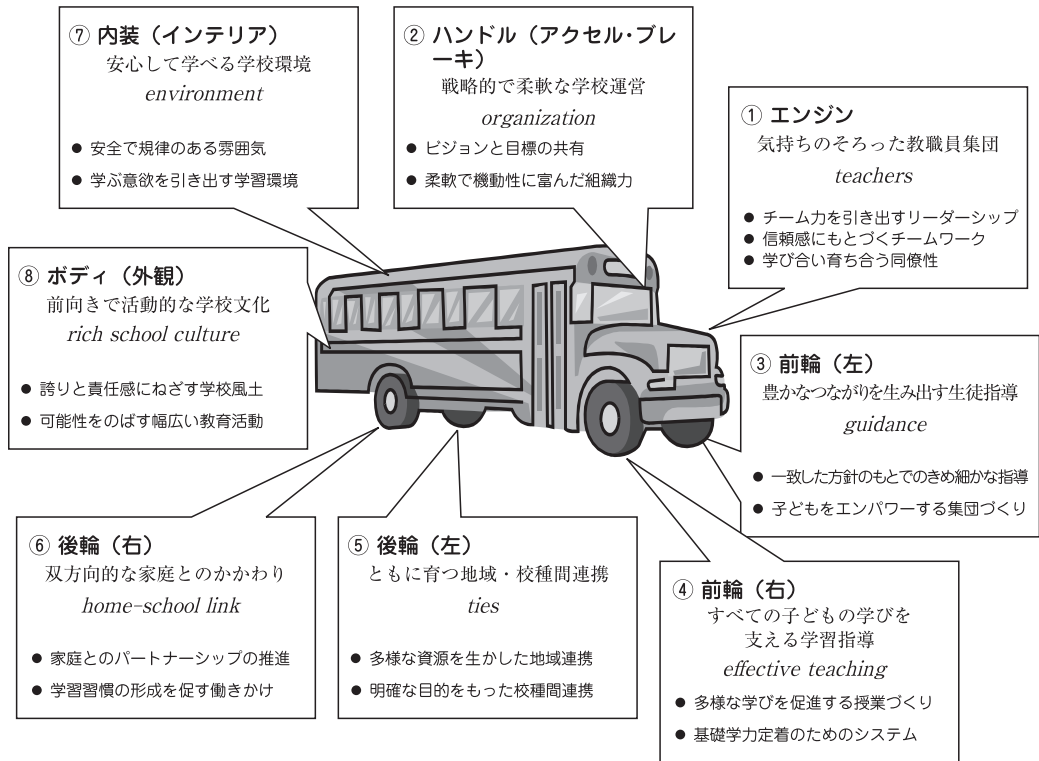


図1 スクールバスモデル

のモデルである。学力調査という量的手法とフィールドワークという質的手法の有機的な組み合わせがなければ、このモデルは決して考案されることはなかったであろうと思われる。

さて、ここでは、この共同研究に関して3つの論点を提示しておきたい。すなわち、第一に、何を学校の「効果」と見るかという〈枠組み〉にかかわる問題、第二に、学校取り組みと「効果」をどのように関連づけるかという〈解釈〉にかかわる問題、そして第

三に、研究成果をどう実践に生かしていくかという研究者〈役割〉にかかわる問題である。

第一の〈枠組み〉の問題。筆者らは、学力テストの点数を引き上げることを学校の「効果」だと捉えた。目には見えない「学力」を操作的に捉えるために「点数」という指標を用いるのは、いわば社会調査の常道であり、私たちにとっては自明のことである。しかしながら、この仕事を通じて痛感したのは、学校現場で奮闘努力している教師たちにとっては、それは必ずしも納得のいくものではない

ということであった。「本当の教育の成果は、テストの点数では測れない」「ペーパーテストの結果だけで学校の善し悪しをうんぬんするのは間違っている」。そのような声を幾度となく聞かされた。

研究者の枠組みの妥当性 (validity)、あるいは適切性 (relevance) が問われているのである。そのような時、筆者は次のように答えるようにしてきた。「ごもっともです。点数だけが教育ではありません。おそらく点数よりも大切なものがあります。ただ、だからと言って、点数がどうでもよいというわけではありません。現に子どもや保護者らは、その点にこだわっています。先生方もそうですね。テストの点数を引き上げることにこだわって学校の成果を議論するのは、決して的外なことだとは思いません」。

第二の〈解釈〉の問題。「効果のある学校」論の視点からデータを分析し、結果を解釈するにつれてだんだんわかってきたのは、「学校の取り組み」と「子どもたちの学力」とは決して直接的には結びつかないということであった。研究に着手した時点で私たちが想定していたのは、「成果が上がっている学校」は「がんばっている学校」に違いないだろうということであった。しかしながら、事実はずうではなかった。各種のクロス分析をしてみると、「がんばっている学校」ほど「成果が上がらない」といった結果が頻繁に見出されたのである。

私たちは当初、そうした結果に戸惑った。しかし、ほどなく事情がわかってきた。

欧米の教育社会学の常識であるが、学校のパフォーマンスは、その学校が立地する地域の社会経済的背景の影響を非常に強く受ける。すなわち、地域の失業率や生活保護率といった数値とその学校のテストの平均点との間に

は、強い相関があるのである。私たちのデータについても、同様のことが言えた。大阪においても、地域の社会経済的状况はテストの点数の出方に大きな影響を与えていたのである。それだけではない。大阪では（おそらくその他の自治体においても）、「がんばっている学校」というのは、往々にして「しんどい」地域にある学校のことが多い。それらの学校では、がんばってはいるものの、なかなかそれが成果には結びつかない。片や、安定した地域に立地する学校では、大した取り組みをしなくても、子どもたちのテストの平均点はかなり高かったりする。そこで、先に見たような「逆説」的な事態が生じるのである。統計用語を使うなら、「学校のがんばり」と「学力」との見かけ上の逆相関は、明らかに「擬似相関」と言ってよいものなのである。

第三の研究者〈役割〉の問題。これは、現在も筆者の頭を悩ませ続けている問題である。私たちの共同研究の成果物である「スクールバスモデル」は、学校現場からの受けが結構よい。すなわち、「わかりやすい」「参考になる」というポジティブな評価を受けることが多いのである。「よりよい学校を作るためのガイドラインに」という実践的な意図をもって作成したものであるから、そうした評価は当然うれしい。しかし、その次がある。「このモデルを参考にして学校づくりに着手したので、協力してほしい。アドバイスしてほしい」。そのような引き合いが、実際かなりたくさん来る。

「現場に役立つ」という意味での研究の実践性を高めたいと考えてきた私たちなので、そのような要望に答えたいという気持ちは強いが、あまりにそちらに行き過ぎるとキリがないというのが、今の率直な気持ちである。最近までの教育社会学は、「現実を理解する」

ことを旨としてきた。しかしながら、「実践」や「臨床」といったキーワードがクローズアップされるにつれ、「現実介入する」教育社会学のあり方が今日模索されようとしている。筆者もそれを試みようとしているのであるが、どこまで介入すればよいのか、そのさじ加減がいまだつかめないでいる。

## 4 施策としての社会調査 ——全国学力テストをめぐって

次に見てみたいのは、子どもたちの学力の実態を把握するために、多額な国家予算をつぎこんで実施されている社会調査の事例である。2007年から実施に移されている全国学力・学習状況調査（以下、「全国学力テスト」）がそれである。初年度における予算は、何と77億円にのぼったという。

この調査では、通常の学力テストにプラスして、生活習慣や学習環境に関する子どもたちおよび学校に対するアンケート調査が付加されている。このスタイルは、PISA調査等に学んだものと言えよう。対象となるのは、全国の小学校6年生と中学校3年生。テストは国語と算数・数学の2教科で、それぞれについて、主として「知識」に関する「A問題」と、主として「活用」に関する「B問題」の2種のテストが子どもたちに課される。2007年度調査に参加した子どもたちの数は、小6が約117万人、中3が約116万人にのぼった。問題と採点基準の作成は文部科学省が行い、調査問題の回収、調査結果の採点・集計等は民間企業（小学校はベネッセ、中学校はNTTデータ）が請け負った。

なぜ、このような全国調査が必要なのか。その実施を正当化する論理としてあげられるのは、以下の4点である。第一に、「全国の学力水準・格差の実態を把握し、成果と課題

を取り出す」こと（【実態把握】の視点）。第二に、「子どもたちの学習の進歩を捉え、指導の改善に役立てる」こと（【教育評価】の視点）。第三に、「教育成果の中心的なものである学力の状況を、市民に目に見える形で伝える」こと（【説明責任】の視点）。第四に、「テスト結果をめぐる競争によって全体の学力向上を図る」こと（【競争主義】の視点）。

筆者は、最初の【実態把握】の視点が最も大切で、3番目の【説明責任】の視点がそれに次ぐ意義を有していると考えている。他方、2番目の【教育評価】の視点を生かすためには全国悉皆調査より、地域や学校の特性に応じた調査が設計されるべきであると思う。また、4番目の【競争主義】の視点は明らかに適切ではなく、有害ですらあると考える（志水、2009）。

さまざまな形で論争を引き起こしている全国学力テストであるが、ここでは、以下の4つの論点にふれておきたい。

第一は、「悉皆調査であるべきかどうか」という、〈調査対象〉の問題である。筆者は、2007年に実施された1回目の全国悉皆調査にはそれなりの意義はあったと思う。かつて1956年から66年にかけて全国学力テストが実施された経緯があるが、その間に悉皆調査がなされたのは、61年から64年にかけての中学校においてのみである。したがって、2007年の調査はじつに43年ぶりの全国悉皆調査だったということになり、子どもたちの学力実態の変化を追うのにこれほど貴重なデータはないと言ってよいし、実際に学力の地域間格差の変貌ぶりを始めとして、いくつかの貴重な結果が見出されたところでもある（志水、2009、27-29頁）。

しかしながら、翌2008年に実施された全国学力テストの結果は、前年のものとほぼ同

じものであった。そうだとしたら、毎年ばく大な予算をかけて悉皆調査をやる意味があるのだろうか。筆者はまったくくないと思う。全国の動向を把握するためであれば、悉皆調査は10年、多くとも5年に一度で十分であろう。数十億にのぼる予算は、少人数指導を行うための人員確保といった、他の用途に割り当てるべきである。

第二に、〈調査方法〉の問題がある。これについては、さまざまな論点があるが、「学力格差の問題に定める」という視点から、1つの問題提起をしておきたい。それは、家庭の情報をどう収集するかという問題である。今回の全国学力テストの場合は、児童生徒用アンケートで、子どもたちの家庭生活であるとか、学習習慣であるとかが尋ねられた。また学校調査においては、個々の学校に関する、就学援助率などのバックグラウンド情報や、学力向上の具体的な取り組みなどについての情報が集められた。それにより、文部科学省がこれまで行ってこなかった類いの分析がなされるようになったことは、評価できる点である。

しかし、それらの分析を通じて、子どもたちの学力の状況を考えるうえで最も重要なのは、おそらく家庭の教育的・文化的環境や子どもへの具体的な働きかけといった要因であるということがわかってきた。プライバシーの問題等クリアしなければならない課題は多いが、保護者の収入・職業・学歴といった基礎的な社会的変数を含めた保護者アンケートの実施の可能性が、今後探られるべきであろう。学力格差の問題に抜本的に取り組むためには、それは不可欠な視点である。

第三に、集められた豊かなデータを、誰が、どのように分析するかという〈結果の分析〉にかかわる問題がある。現状では、2度の全

国悉皆調査によって収集されたばく大な量のデータは、いわば「宝の持ち腐れ」状態となっている。端的に、調査主体である文部科学省には分析に必要なマンパワーと時間的余裕が不足しているようである。データは、都道府県・市町村にフィードバックされ、各自治体で主体的に分析を行い、改善のための対策をとるべしというスタンスがとられている。「丸投げ」とは言わないが、責任はそれぞれの自治体がとるべしと言わんばかりである。もちろん、自治体の中には精力的に分析を行い、着実な施策立案に結びつけているところも多いのである。

本来であれば、文部科学省は、研究者や政策担当者からなる分析チームを立ち上げ、系統のかつ長期的な視点からのデータ分析を行うべきであろう。しかしながら、現状ではそうした動きはない。その代わりとして、文部科学省が今年度（2008年度）立ち上げたのは、研究者グループへの研究委託事業である。筆者は、そのうちの1つ、お茶の水女子大学の耳塚寛明教授を代表者とするグループに現在参加している。そのグループの使命は、「格差問題に対応した学力調査の設計」を開発することにあり、2009年には何らかの研究結果を発表する予定である。

第四の論点として、「調査結果をどう公開するか」という〈結果の公表〉の問題を指摘しておきたい。2008年秋には、この点をめぐって、鳥取県や大阪府などいくつかの自治体の動きが大々的にメディアに取り上げられた。とりわけ大阪では、多くの市町村教育委員会の反対があったものの、結局橋下知事が、市町村別の平均正答率の「一部開示」を行った。「一部開示」と形容してはいるが、多くの市町村の数値の一覧表が新聞紙上等に掲載され、大きな波紋を呼んだ。知事サイドは、

【競争主義】の視点に立って開示を推し進めたわけであるが、市町村サイドの抵抗には少なからざるものがあつた。自治体別平均正答率の公表は過度の競争や序列化につながる、という危惧がその背景にある。

前節でふれた「効果のある学校」論的な観点からしても、単純な平均点による順位つけはのぞましくないものだと言える。というのも、先にもふれたように、子どもたちの学力には「家庭の力」（地域の社会経済的状況と強く関連する）と「学校の力」（学力向上の取り組みに代表される）との両者がかかわっているからである。あたかも「教育の質」や「学校の優秀さ」を示すように見える平均点の高さには、いつも「家庭の力」の高低が暗数として隠されている。紙幅の関係で詳しく述べることはできないが、欧米の研究では、その点を考慮に入れた理論的工夫（「付加価値」方式などと呼ばれる）や方法論的洗練（「マルチレベル分析」が代表的である）がなされつつある。日本でも、それらの積極的な受容がのぞまれるところである。

## 5 まとめ

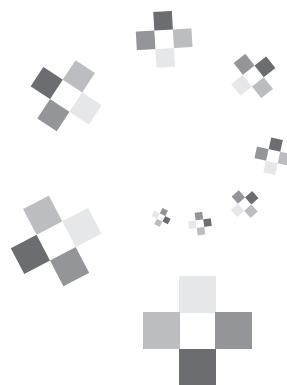
日本では、1950年代から60年代にかけて全国的な学力テストが実施された経緯があるが、それ以降40年以上の長きにわたって、同程度に大規模な全国調査は行われてこなかった。と同時に、「学力テスト」はもっぱら心理学者や教育学者の占有物として扱われ、社会学者の出る幕はほとんどなかった。欧米の教育社会学では、社会集団・社会階層と学力という主題は、メイン中のメインと言ってもよい研究テーマであり続けてきたが、日本の教育社会学においては、一部の例外（代表的なものは、関西地方で行われてきた同和地区児

童生徒の低学力の実態を把握する調査）を除いて、この数十年、表舞台に現れてくることはなかった。知的怠慢と言われても仕方のない状況である。

今や日本は、「学力新時代」に突入した感がある。「今こそ私たちの出番である」と、筆者は強く感じている。社会調査の精神と技術を有した私たちこそが、喚起力・洞察力を有する研究成果を発表し、学力をめぐる社会的議論を適切に導いていかなければならない。学力問題を解く鍵は、誠実な学問的探究の先に見出すことができるだろう。

### 文献

- 市川伸一，2002，『学力低下論争』筑摩書房（新書）。  
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。  
苅谷剛彦・志水宏吉・清水陸美・諸田裕子，2002，『調査報告「学力低下」の実態』岩波ブックレット No. 578，岩波書店。  
鍋島祥郎，2003，『効果のある学校——学力不平等を乗り越える教育』部落解放・人権研究所。  
志水宏吉，2003，『公立小学校の挑戦——「力のある学校」とはなにか』岩波ブックレット No. 611，岩波書店。  
———，2005，『学力を育てる』岩波書店（新書）。  
———，2009，『全国学力テスト——その功罪を問う』岩波ブックレット No. 747，岩波書店。  
山内乾史・原清治編，2006，『学力問題・ゆとり教育』日本図書センター。







## 調査フィールドとしての学校

——アクセスの困難さとアクションリサーチへの期待——



日本では、1980年代から多くの研究者が学校でフィールドワークを実施し成果を挙げてきた。だが今日、学校へのアクセスは、かなりの困難が生じている。本稿では、その背景を探る中から、調査フィールドとしての学校はどのような特性を有しているのか、今日そこにいかなる変化が生じつつあるかについて考える。また、その変化に対応する中で、学校をフィールドとする調査は、臨床性が強く求められるようになり、アクションリサーチが強調されるようになった。その経緯と、筆者自身の取り組みについても報告する。

はじめに次の2点を確認しておきたい。第一に学校とは、ある教育目標に基づいて編成されたフォーマルな組織である。ただし、いくつかの点で他の組織とは異なる特性を有している。第二に、学校をフィールドとする研究者には教育社会学を専攻する者が多いが、天野（1990）が言うように、教育社会学者は社会学と教育学の間であって境界人性を有している。研究者によりそのいずれにより多くのアイデンティティを感じるかは異なっており、職務上も教育学部で教員養成に携わっている者がいる一方で、社会学系のポストに就いている研究者もいる。

調査のフィールドとしての学校のあり方やそこへのアクセスは、学校組織の特性と、フィールドワーカー自身のアイデンティティの

酒井 朗（大妻女子大学家政学部児童学科教授）

あり方の双方に規定される。本稿ではこのことを念頭に検討を進めていく。

### 1 学校をフィールドとする調査への関心の高まり

学校の内部を明らかにしようとする問題関心は、欧米では1960年代末から台頭し、日本もそれより約10年遅れて急速に高まった。

学校をフィールドとする調査研究の1つの流れは、学校の選抜機能を解明するために、学校内部の選抜過程に光を当てようとするものである。イギリスでは1960年代後半から70年にかけて、Hargreaves（1967）やLacey（1970）らがこうした研究に着手した。また、日本でもそれにヒントを得た研究が80年ごろから開始された（耳塚，1980）。いわゆるインプット-アウトプット研究に対してスループットを解明せよと言われたのがこの流れの研究の合い言葉であった。

またM. ヤングの提唱した新しい教育社会学の流れは、再生産や学校の社会的選抜機能という教育社会学の基本問題に関心をもちながらも、学校知や教師の判断枠に対する関心を強めた。1971年に刊行された教育社会学の代表的リーディングスのタイトル、『知識とコントロール』は、当時の研究関心のあり方を端的に示している。イントロダクションでは、社会学者は学校で何が教えられるか自

体を批判的に分析し、知識の編成の問題と支配の問題、つまり「知識とコントロール」がどのように関連しているかを明らかにしなければならないと唱えられた (Young ed., 1971)。

また、エスノメソドロジストも制度的場面の1つとして学校内部を丹念に観察した。代表的な Mehan (1979) の研究は教室における教師生徒間の会話の特徴的なパターン (教師の発問 I-生徒の応答 R-教師による評価 E) を指摘し、今や学校研究における基本仮説となっている。

## 2 学校へのアクセスを阻害する要因

このように日本でもさまざまな問題関心に基づいて学校をフィールドとする研究が増えてきた。それらの諸成果を見ると、多くの学校がフィールドワーカーを積極的に受け入れてきたかのように映る。だが、成果として外部に発表されるプロジェクトの背後には、多くの挫折したプロジェクトがあり、またフィールドに入る前に頓挫した計画もある。学校は、フィールドワーカーを必ずしも歓迎してくれるわけではない。彼らが学校に入りこみ、そこで調査を継続させるには相当の困難が伴ったのであり、しかも近年、その度合いはますます強まっているように思われる。

この種の問題は、フィールドの成果を著した論文には書かれず、方法論の教科書でもあまり触れられることがない。しかし、今後学校をフィールドとする調査がさらに発展するうえでは、本誌のような社会調査関係者間の情報交換の場を通じて、「論文では語られないこと」を議論することも必要だと思われる。以下の検討は、もっぱら筆者自身の経験に頼る部分が多いが、あえてそれをたたき台として報告したい。

日本の学校では、大きく分けて次の6つの要因が、フィールドワーカーのアクセスを阻害していると思われる。

第一に、よく言われるように日本の学校は閉鎖的であり、部外者が入ることはそもそも難しい。イギリスのコンプリヘンシブ・スクールで参与観察調査を実施した志水宏吉は、調査を依頼した学校がすべて氏の申し出を受け入れたことに驚き、日本の学校の閉鎖性を際立たせている。氏が同国で得た説明では、学校にはアカウントビリティがあり、外部者からの見学依頼を断ることは基本的にはできないとのことであった。

第二に、最近の傾向として、個人情報保護法が大きな足かせになっている。とくに神経を使うのはビデオ録画である。教室での観察や会話の採取にはビデオを使うことが多い。しかし、児童生徒を撮影することについて、個人情報保護を理由に学校側から断られるケースがある。筆者が現在観察している幼稚園は、日ごろから研究熱心で、保護者に対し研究を目的としたビデオ撮影について了解を得ている。このため、筆者らがビデオを持ち込もうとした時も比較的容易に承諾された。反対にある養護学校での活動を記録したビデオは、個人情報保護の観点で部外者の閲覧は強く制限された。

第三に教師の多忙化が年々深刻化する中で、教員にこれ以上の負担をかけたくないという理由で管理職から調査を断られることが多い。ベネッセ教育研究開発センターが小中学校の教員約4000名を対象に実施した第4回学習指導基本調査(2007年8月~9月実施)によれば、大半の教員が「教材準備の時間が十分にとれない」「作成しなければならない事務書類が多い」「休日出勤や残業が多い」と回答した。このような中で学校では、「外部の仕

事」はできるだけ引き受けないようにセーブすることが多く、アンケート調査の依頼も拒否にあうことがある。フィールドワークで学校に調査に入るとは教職員にはそれ以上に負担感が大きいので、調査対象校の管理職や教師となんらかのコンネクションがないかぎり、調査が許可されることは難しくなっている。

第四に挙げられるのは、学校組織の構造上の問題である。学校組織論では、学校は官僚制の発達した他のフォーマル組織と比べて各部署間の連結が緩やかだと指摘されている(Weick, 1976)。企業や官公庁であれば上部の決定に部下は従うものと期待される。だが、学校では必ずしもそのように進まないことも多い。管理職が了承しても、各学年の主任や担任教師からクレームが付いて調査が頓挫してしまうこともある。

筆者は、多くの生徒が中学に上がると学校不適應を生じさせる社会的要因を探るため、数年前にある公立中学校に、次年度に入学する生徒を最初の数ヶ月間観察させてほしいと願い出たことがある。当初、学校長は調査研究の意義を理解し、部活動や学校行事の時間などであれば学校に来て調査することはかまわないと承諾した。しかし、その後担任団から、入学早々に校内の活動を細かく観察されると教育活動に支障をきたすという理由で断られた。

緩やかに連結された組織構造では、管理職だけではなく、関係するすべてのメンバーに調査のねらいや調査の方法について了解を得なければならぬ。フィールドワークの教科書には、フィールドへのアクセスには、ゲートキーパー(門番)に了承を得ることが重要だと書かれている。だが、学校には何人もの「門番」がいるように感じられる。ただし、このことは学校差が大きく、管理職が承諾す

れば比較的すんなり受け入れられるケースもある。ある意味では学校へのアクセスの段階で、学校の組織的特性や、対象校の学校文化の一端をうかがい知ることができるとも言える。

第五の要因は教員異動の多さである。周知のことかもしれないが、近年は公立学校教員の異動のインターバルがたいへん短くなっている。東京都では現任教員において引き続き3年以上勤務する者は異動の対象者となる。どの学校でも年度の変り目に大量の異動が生じるため、年度替わりの前に次年度の調査について管理職と話し合うことは難しく、また、調査継続中でも管理職が異動し、継続が難しくなることもある。フィールドワークにおけるフィールドとは、もともとある場所(土地)への参入を意味したもので、ゲートキーパーとはその土地の有力者であることが多い。そこに含意される安定性とは対照的に、日本の公立学校はフィールドとしてははなはだ不安定であり、継続的な調査はかなりの困難を伴う。

第六の要因は、社会学的調査の本質に関わる問題である。そもそも社会学的な研究とは、エスノメソドロジーがそうであるように、人々が自明としている日常を異化し、それがいかにして成立しているのか=社会はいかにして可能かを問うものである。また、機能的な観点から、その営為に隠された機能を暴こうとする志向性も強い。

学校をフィールドとする場合も社会学的視点を携えてフィールドワーカーは入り込む。そして、その結果得られた知見はしばしば教師の常識的思考を打ち破り、また彼らの行為の隠れた帰結を暴き出す。しかし、この種の知見について、学校側に報告した際にしばしば遭遇するのは、教師の当惑であり、「そん

な目でわれわれを見ていたのか」という批判である。

エスノメソドロジーにおける違背実験がそうであるように、自明な現実を崩された人々は怒りだすことが多い。もちろんそのような状況は他のフィールドでもしばしば生じていることと思われる。だが、学校は何らかの教育理念に支えられている分、それを揺さぶるような指摘は、「われわれが子どものためだと思って日々がんばってきたことを、あなた方はそのように批判するのか」という感情的な非難を招きやすい。子細を述べることは出来ないが、ある学校を調査した際に、対象となった教師からの依頼で、実際にわれわれが観察から見出したことを報告する機会をもった際に、われわれはそのようなりアクションを受けたことがある。

先に述べたように、多くの学校で多忙化が深刻化しているが、それをかろうじて支えているのが教師の教育理念であったり、子どもへの思いであったりする。それだけに、フィールドでの成果を著そうとすれば、こうした教員のがんばりを否定するような記載はしにくいものである。

フィールドワークを書く際に、学校側の思いを汲んで配慮を加えることを、志水は「筆を鈍らせる」という表現で説明している。「エスノグラフィックな研究の再帰的性格を考えると、筆が鈍るのはけだし当然と言えるかもしれない」（志水、1998：21）と氏は述べているが、学校をフィールドとする研究は、そうした配慮を加えつつ著されることが多いと言えよう。

### 3 「臨床」研究への志向性

学校へのアクセスが困難になる一方で、学

校現場の課題に沿った研究成果の発信が強く求められている。教育学の領域では、そのことが「臨床」をキーワードに求められている（酒井、2002、2004）。臨床とは、問題が生じているその場において、その問題の解決にむけて取り組む姿勢をさす。研究者がその解決を志向して学校現場に深く分け入り、問題の構図や背景を丹念に明らかにしていくことが期待されているのである。このように整理すると、学校をフィールドとした調査は、一方で大きな困難を生じさせながらも、他方ではある志向性を有した研究は、強く期待されることが分かる。

学校をフィールドとする研究者はこのような事態に直面しており、積極的にせよ消極的にせよ「臨床」ということを謳わずには、調査を実施することは困難な状況にある。しかし、社会学者はカウンセラーのように児童生徒をクライアントとして、彼らの治療に当たるという関係を取るものではない。それゆえ、学校をフィールドとする研究者は、いかにしてこうした現場の期待に添いつつも、社会的なスタンスを維持して調査を進めるかという課題に取り組まざるをえなくなっているのである。

この課題に関連して、秋葉と広川（2005）は、臨床的研究における当事者、研究者、読者の関係をタイプ1～3の3タイプに分けて整理している。カウンセラーのような関わりはタイプ2に挙げられている。

フィールドワークは、タイプ1とタイプ3として挙げられている。タイプ1は、当事者が抱えている問題の解決に貢献しようという志向性を強く持ち、現場に出向きエスノグラフィックな調査を進めるものである。これに対して、秋葉らが独自に提唱しているのがタイプ3で、学校現場で得られた会話のトラン

スクリプトを、当事者、研究者、読者が再演する作業をコアに据えて、それぞれの観点からその現場での「アクチュアリティ」（当事者が、いま、この場面で何をどのように気にかけてつつコミュニケーションしているのかという、リアリティのもつ当事者性）を追体験させるという手法を取るものである。彼らによれば、タイプ1、タイプ2がその場で生じている問題の解決になんらかの形で対応することを目指すのに対して、タイプ3は当事者や読者にその場で何が生じているのかを客観視させることを通じて問題への対応を再考させる効果を期待する。

また、古賀（2008）は従来のフィールドワークが研究者の観点でまとめてきたことを批判し、構築主義的エスノグラフィーの可能性を追求することを提唱している。構築主義的エスノグラフィーでは、組織内で立場や利害の異なる人々の声を聞き取り、その多声性を表象することが重要だと指摘される。そして、「収集整理されたテキストの分析は、優位な一致する解釈へと向かうのではなく、可能な解釈の構造へと向かっていき、二項対立的な理解の図式そのものを表象し、さらには分析されたテキストを通して読者へと再演されることになる」（古賀、2008：171）。氏は、フィールドにいるさまざまな他者の声を尊重するという姿勢で現場との接点をさぐるようとしている。

## 4 アクションリサーチへの期待の高まり

学校調査をめぐる現状へのもう1つの対応は、学校が抱える問題に焦点を当て、その解決にあたることを含めて研究として取り組もうとするものであり、アクションリサーチと呼ばれることが多い。秋葉らは、これをタイ

プ1に含めているが、従来のエスノグラフィーとは異なり、フィールドで生じている問題の解決に、より積極的にあたろうとする点に特徴があり、近年期待が高まっている。

アクションリサーチとは、「社会的な課題や問題に応じて実施され、その成果がその状況を改善するために利用されるような、社会的に有益でかつ理論的にも意義がある研究の総称」だとされている（鹿毛、2002：160）。ただし、社会心理学者のレヴィン（1954）が提唱した元々のアクションリサーチは、厳密な実験研究とフィールドとを連結しようとする社会工学的な研究方法で、「計画」「実行」「査定」をスパイラルで反復し実践を発展させていこうとするものであった。

かりにこの延長線上でアクションリサーチを捉えれば、フィールドワーカーは、さらに新たな困難に直面せざるをえない。なぜなら調査フィールドとして学校に入る場合、学校とは理解を試みる対象であり、そこにいる教員や生徒は重要なインフォーマントであって、いわばフィールドワーカーにとっての教師である。これに対して、アクションリサーチャーとして現場に入る場合は、対象者に評価的に関わることもありうる。フィールドワーカーがアクションリサーチを行おうとすると、自分と相手との関係をどう構築しながら、実践に関与していくのかという課題に直面することとなる。

また、このことは、研究者としてのアイデンティティという、冒頭で述べた問題にもつながっていく。社会学者として問題を異質化整理する立場に立てば、学校はフィールドワークの場として存在する。しかし、教育学者というアイデンティティに立てば、むしろそれが抱える諸課題の解決にあたり、教育実践に対して方向性を示すことにウエイトがかか

る。教育学系の学部学科や教職課程を担当する者は、そうした役割期待に応じざるをえない。したがって、その役割期待に応えつつ、一方でフィールドワーカーとしての立ち位置を守っていかなければならなくなる。

こうした役割葛藤を内面に孕みつつ、その中でアクションリサーチを遂行しようとする場合、選択しうるのは参加型アクションリサーチの形態であろう。そこでは、研究者は研究プロセスをともにする共同研究者を意味するものとなり、現場とは縦の関係ではなく、むしろ参加者の一員であるとみなされる。鹿毛(2002)は現場教師との対等な関係を強調する場合の研究者のスタンスをコ・アクションリサーチャーと呼んだ。すなわち共働するアクションリサーチャーである。

これまで筆者自身はアクションリサーチとして2つのプロジェクトに取り組んできた。1つは、幼稚園と小学校の連携についてであり、もう1つは、高校生の進路選択支援活動に関する研究である。これらにおいて、コ・アクションリサーチャーとして筆者が果たした役割は、基本的には教師の指導観(保育間)や生徒の進路意識を理解し、それを整理し概念化し、改善策や支援の方策を検討するという、フィールドワーカーの延長線上にある作業を引き受けることであった。筆者はプロジェクトを推進する教師や関係者とのやりとりの中で、改善策や対応策を検討するとともに、研究的にはその過程を整理して記述し、それを通して学校や教師、子どもの特徴について理解を深めることを課題としてきた。

幼小連携のプロジェクトでは、幼稚園教諭と小学校教諭とともに、連携を深めるためのカリキュラム開発というプロジェクトに関与した(酒井ほか, 2004)。筆者が果たしうる役割は双方の言葉を相手方に翻訳し、お互いの

共通理解を深めていくというものであった。異なる校種間の連携を図るうえでの「異文化理解」の橋渡し役として機能したわけである。

また、ある商業高校で実施した高校生の進路選択を支援するプロジェクトでは、生徒の進路意識をライフストーリー論の観点から分析し彼らの理解を深めるとともに、学生ボランティアを組織し、その学校の教員と協働して、生徒の進路選択を支援する活動に従事した(酒井編, 2007)。そこでわれわれが重視したのは、高校生がボランティアと場を共有する中で、大学に関する情報を得て、進学に向けて動機付けられることであった。進学者がけっして多くはないその学校の中に、われわれは、進学にむけて生徒が意識を高めていくために特別な場を設けたわけである。筆者はこの取り組みの中でも、フィールドワーカーとして解釈的に行為者の世界を追体験しつつ、彼らを取り巻く環境の調整に当たったのであった。

## 5 いくつかの課題

学校でアクションリサーチを進めるうえでしばしば指摘されるのは、教師自身がリサーチャーとして関与することである。2008年度から開設された教職大学院では、現職教員を中核的リーダーに養成するために、フィールドワークやアクションリサーチを授業に取り込むことが期待されており、今後「リサーチャーとしての教師」が増えていくものと予想される。学校における質的調査が拡充するには、こうした新しいタイプのリサーチャーの参入が期待される。また、彼らの養成に資するカリキュラムについてもさらに検討が求められる。

ただし、それと同時に必要なのは、アクシ

ョンリサーチ以外の研究スタイルも十分に尊重されるとい研究風土である。研究の実践への貢献が強く望まれているのだが、それとともに科学としての批判性をいかに担保するかという課題を追求していくことが必要である。

また、アクションリサーチへの傾斜が、フィールドワークを通じて開発されてきた各種の分析技法を軽視することになってはならない。コーディング、カテゴリー生成などの基本的技法や、会話分析、ライフストーリー分析などの分析手法をアクションリサーチとどうつなげていくのかも検討する必要がある。

フィールドとしての学校が抱える課題はすでに指摘したとおりであるが、やはり最大の課題は、学校の閉鎖性である。組織の閉鎖性は学校教育の意義や課題を理解し、改善していくことを妨げてしまうことにもなる。

以上、さまざまな課題が今後対応されなくてはならないが、それらの取り組みを通じて、学校現場に対し意味のある調査研究がいっそう増えることを期待している。

#### 注

- ・1 日本心理学会第69回大会公開シンポジウム「教育のアカウントビリティを問う」2005年9月11日(日)での志水宏吉氏の報告「教育社会学の立場から」より。

#### 文献

- 秋葉昌樹・広川義哲, 2005, 「演劇的問題表象のメソドロジー序説——リアリティとアクチュアリティの往還と教育研究における臨床性」『龍谷大学教育学会紀要』4: 17-35。
- 天野郁夫, 1990, 「辺境性と境界人性」『教育社会学研究』47: 89-94。
- Hargreaves, D. H., 1967, *Social Relations in a Secondary School*, Routledge & Kegan Paul.
- 鹿毛雅治, 2002, 「フィールドに関わる『研究者/私』——実践心理学の可能性」下山晴彦・子安増夫編『心理学の新しいかたち——方法への意識』誠信書房, 132-72。

古賀正義, 2008, 「構築主義的なエスノグラフィーを实践する」北澤毅・古賀正義編『質的調査法を学ぶ人のために』世界思想社, 153-77。

Lacey, C., 1970, *Hightown Grammar: The School as a Social System*, Manchester: Manchester University Press.

レヴィン, K., 1954, 『社会的葛藤の解決——グループダイナミクス論文集』末永俊郎訳, 創元新社。(K. Lewin, *Resolving Social Conflict*, Harper & Row, 1948.)

Mehan, H., 1979, *Learning Lessons: Social Organization in the Classroom*, Cambridge, Mass.: Harvard University Press.

耳塚寛明, 1980, 「生徒文化の分化に関する研究」『教育社会学研究』35: 111-22。

酒井朗, 2002, 「臨床教育学構想の批判的検討とエスノグラフィーの可能性——『新しい教育学の創造』と『問題への対処』をいかにして同時達成するか」『教育学研究』69(3): 322-32。

——, 2004, 「教育臨床の社会学——特集にあたって」『教育社会学研究』74: 5-20。

——編, 2007, 『進学支援の教育臨床社会学——商業高校におけるアクションリサーチ』勁草書房。

——・藤江康彦・小高さほみ・金田裕子, 2004, 「幼小連携におけるカリキュラムの開発に関するアクションリサーチ」お茶の水女子大学21世紀COEプログラム『誕生から死までの人間発達科学——家庭・学校・地域における発達危機の診断と臨床支援』, 51-70。

志水宏吉, 1998, 「教育研究におけるエスノグラフィーの可能性——『臨床の知』の生成に向けて」同編『教育のエスノグラフィー——学校現場のいま』嵯峨野書院, 1-28。

Weick, K. E., 1976, "Educational Organizations as Loosely Coupled Systems," *Administrative Science Quarterly*, 21(1): 1-19。

Young, M. F. D., ed., 1971, *Knowledge and control: New Directions for the Sociology of Education*, London: Collier-Macmillan.





# 3

## 親子調査における親欠票の原因

——仙台高校生調査データを用いた分析——



神林 博史 (東北学院大学教養学部准教授)・片瀬 一男 (東北学院大学教養学部教授)

近年、面接調査に関しては、とくに大都市における調査拒否と一時不在の増加によって回収率が低下していることが指摘されているが(坂元, 2001), 回収率の低下は学校における自記式の配票調査についても生じている。

本稿では、1986年度から2007年度まで仙台圏で継続的に実施されてきた「教育と社会に対する高校生の意識」調査(以下、「仙台高校生調査」と略す)における親欠票の分析を行う。本調査は学校を通じた親子調査であるため、どのような親が調査に回答しないのかについて、学校側あるいは高校生側の情報を用いることで、ある程度まで調べることができる。本調査の場合、高校生に比べて親、母親に比べて父親の回収率が低い。欠票分析に限らず、あるグループの特徴を分析するためには、当該グループの標本数は多い方が良い。そこで本稿では、欠票の多い(=回収率の悪い)父親について特に詳しい分析を行う。そして親欠票の分析から、学校を通じた親子調査における親の回収率低下の原因と、それが引き起こすと考えられる偏りを検討し、そこから見えてくる家族、教育、社会の変化について考察する。

### 1 仙台高校生調査の概要

仙台高校生調査は、仙台圏の高校を調査対

象にして、高校生が日本社会や学校教育に対してどのような意識をもっているのか、またそうした意識の形成に家族や学校がどのような影響をおよぼしているのかを解明するため、1986年度、94年度、99年度、2003年度、07年度の5回にわたって行われてきた(調査主体は、東北大学教育文化研究会)。この調査の特長の1つは、高校生とその父母を対象者とし、親子・夫婦(父母)をペアとして関連づけとともに、高校生と親の意識の趨勢的变化を分析できるように設計されていることである。

この調査の対象者は、層化三段抽出によって抽出されてきた。すなわち、まず仙台圏の高校を高校種別や別学・共学、私立・公立によって層化したうえで学校を抽出した。次に当該高校とも相談のうえ、対象クラス(原則として2年生)を選び出した。最後に、そのクラスの生徒と両親を調査対象者とし、依頼状で趣旨等を説明して協力を得ている。調査方法は、原則として高校生は自記式集合調査、両親は自記式配票調査である。すなわち、高校生にはホームルームや授業等の時間を利用して担当の教員から調査票を配布し、その場で記入してもらった。父親・母親については、生徒を通じて、調査の依頼状、調査票(父親用、母親用の2種類)、調査票回収用の封筒(個人情報保護のため記入済み調査票を密封する)の3点を家庭に持ち帰ってもらい、記入して



表1 仙台高校生調査の回収率

実施時期	計画標本数	対象者					
		高校生		父親		母親	
		有効回収票	回収率	有効回収票	回収率	有効回収票	回収率
1986年度	1312	1225	93.4%	984	75.0%	1108	84.5%
1994年度	1711	1542	90.1%	1203	70.3%	1351	79.0%
1999年度	1409	1259	89.4%	993	66.2%	1097	77.9%
2003年度	1280	1113	87.0%	719	56.2%	877	68.5%
2007年度	1555	1231	79.4%	934	60.2%	1157	74.6%

注) 回収率の計算には高専の対象者が入っている。

もらった。記入済みの親調査票は、生徒をつうじて回収した。なお、同一の対象番号をつけたひと組の調査票（高校生票、父親票、母親票）を高校生に渡すことによって、親子の対応を確保した。この際、回答者の匿名性は確保されている。また調査に関する問い合わせは、学校ではなく研究会にするよう連絡先も明記してあるが、父母からの問い合わせほとんどない。高校によっては、父母宛に校長名で依頼状を入れてくれる場合もあり、調査環境としてはかなり良好な状況にある。

調査終了後にはすべての調査協力者に礼状を配布するとともに、年度末には10ページ程度の「調査結果の速報」を高校生をつうじて家庭に配布している。また、学校にはその後「学校別集計」（すべての間について当該校と対象校全体の比較をしたもの）を届け、校内の研修等での活用をお願いしている。さらに「最終報告書」（内容は科研費報告書と同じもの）は、調査対象校だけでなく教育行政関係機関・宮城県内の全高校に届けるとともに、宮城県高等学校社会科教育研究会の研修会で報告する機会もいただいている。

このように本調査は、学校現場と密接かつ頻繁な交渉を長期にわたって続けており、その結果、高校生票・父母票とも高い回収率を維持することができた。しかし、近年の回収率低下の流れは仙台高校生調査にも及んでいる。表1に示したように、1986年調査に比

べると、2007年調査では高校生票14ポイント、また父親票で15ポイント、母親票では10ポイントほどの回収率の低下がみられる（ただし、注目すべきは最新の2007年調査では親の回収率の低下傾向に歯止めがかかったことである）。そして、この回収率の低下した層が特定の層に偏っているとすれば、調査結果に何らかの偏りが生じている可能性も否定できない。

ただし、仙台高校生調査では、親票が欠票となっている場合でも、高校生票は有効であることが多い。したがって、高校生の在籍する学校の情報や高校生の提供した親に関する情報を用いて、親票の欠票の背景を推測することができる。そこで、次節からは仙台高校生調査のデータをもとに親子調査における欠票を規定する要因について考察していきたい。

## 2 親子調査における非協力の原因

仙台高校生調査の調査法の場合、高校生の非協力者については、対象者に協力する意思がなかったためと単純に解釈することができる。しかし両親の非協力については、家族構成の問題あるいは調査票が高校生経由で配布されることによって生じる問題など、複数の原因が考えられる。親の欠票が生じる原因は、以下のように分類できるだろう。

(1) 親が不在のため欠票になる場合（構造的欠票）

(1-1) 離死別により，親が不在のため生じる欠票

(1-2) 長期出張，単身赴任，入院等により，親が一時的に不在のため生じる欠票

(2) 親は不在ではないが欠票になる場合（非構造的欠票）

(2-1) 高校生が親に調査票を渡していないために生じる欠票

(2-2) 調査票は親に届いているが，親に協力の意思がないため生じる欠票

残念ながら，データの制約からこれらの要因をすべて検討することはできない。とはいえ，構造的欠票と非構造的欠票の基本的な部分については，ある程度の分析が可能である。

### 3 親票の構造的欠票と非構造的欠票

#### ❖ 親不在の影響

仙台高校生調査では，高校生票で家族構成を質問しているため，父親，母親のそれぞれについて構造的欠票（親不在による欠票）がどの程度生じているかについて把握が可能である。ただし，質問文の関係から，親の離死別と一時不在を厳密に区別することはできない<sup>4</sup>。

まず，親不在の高校生がどの程度いるか確認しよう。表2の左半分に，各調査時点の親不在率<sup>5</sup>を示したものである。

親不在率は，父親については近年になるほど上昇傾向，母親についてはほぼ一定である。したがって，父親の回収率低下に関しては，親の構造的不在の影響が無視できないことになる。このような親不在率，特に父親の不在率の増加は，離婚率や離婚に伴って生じる母

表2 高校生の親不在率と欠票における親不在率

調査年	高校生の親不在率 (%)		親欠票における親不在率 (%)	
	父親	母親	父親	母親
1986年	5.5	1.2	20.8	10.1
1994年	5.2	1.6	15.0	6.8
1999年	5.2	0.9	12.8	3.2
2003年	6.6	1.8	12.8	5.1
2007年	8.0	1.2	15.7	3.6

子世帯の動向に関する統計が示す傾向とも一致している<sup>6</sup>。

次に，表2の右半分には，親欠票における親不在率を示した。これは親の欠票（非協力<sup>7</sup>）全体に対して親不在が占める比率である。この数値が高いほど，親欠票が構造的な理由によって生じていることを意味する。

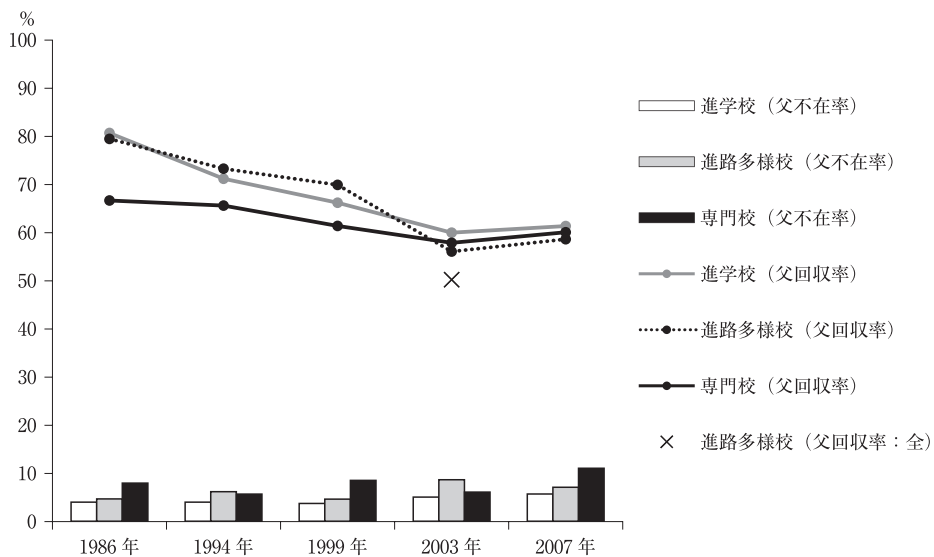
親欠票に占める親不在率は基本的に低下傾向だが，父親が2007年調査でわずかに上昇している一方，母親は大きく変化していない。このことは，2007年調査における親回収率の回復と併せて考えると興味深い。

#### ❖ 高校種別との関係

次に，高校種別と親欠票の関係を分析しよう。ここでの高校種別は，海野・片瀬編（2008）に準じ，(1)進学校（専門学校進学率と就職率の合計が15%未満の普通科高校），(2)進路多様校（専門学校進学率と就職率の合計が15%以上の普通科高校），(3)専門高校（職業科高校），の3カテゴリーとした。

よく知られているように，高校種別と親の社会階層（親職業，学歴）の間にはある程度の関連が見られる。高校種別で比較することによって，欠票に対する親の社会階層，あるいはそこから派生する意識の影響を推測することができるだろう。

図1は，高校種別ごとの父親票回収率（折れ線グラフ），および父親不在率（棒グラフ）



注) 2003年調査では、進路多様校の中に親票の回収率が極端に低い高校が1校存在したため、その高校を除外した値(折線)と、その高校を含めた値(x)を示した。

図1 父親票有効回収率(高校種別別)

を示したものである。回収率について見ると、1986年調査では専門高校とそれ以外の高校の間に10ポイント以上の開きがあったが、近年になるにつれて高校種別による差が消失し、2003年および2007年では、高校種別の差はほとんど消失している(母親の回収率にも同様の傾向が存在する)。

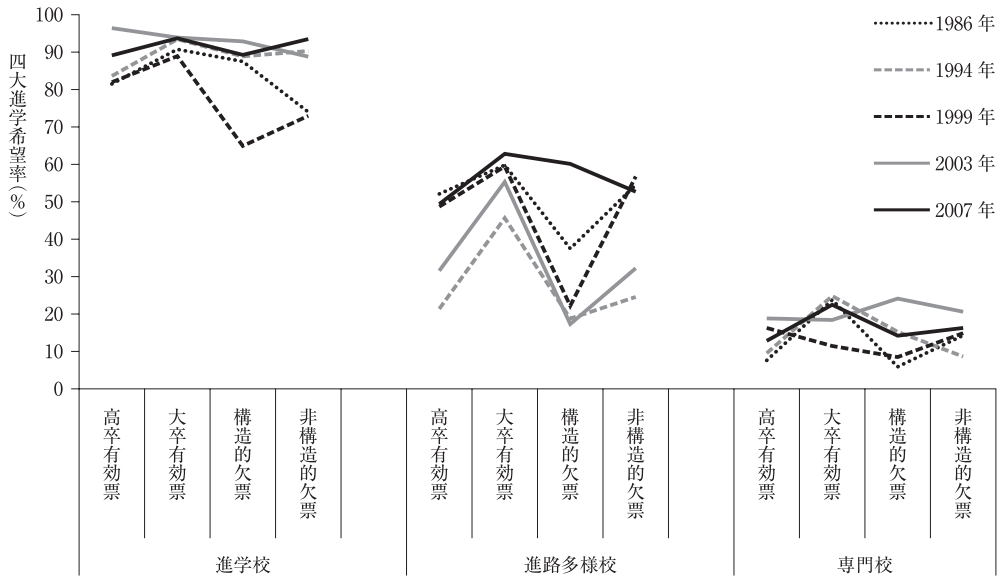
ここから、1990年代までの専門高校の親には、調査非協力を生み出すメカニズムが他の高校よりも強く働いていたか、あるいはこの層の親に特有の傾向が存在していた、と考えることができる。たとえば、長時間労働あるいは不規則な勤務時間のため、回答する時間がなかったり高校生から調査票を受け取れない親が多かった(一般に、専門高校の父親には下層ホワイトやブルーカラー労働者が多い)、子どもの教育にそれほど熱心ではないため「教育と社会」と銘打った調査には協力しない親が多かった、等である。ただし、専門高校の親不在率は他の高校に比べるとやや高いので、この点を差し引く必要がある。

2000年代に入ってから、学校差がほとんど消失したということは、(1)90年代までの専門高校に存在した非協力発生のメカニズムが進路多様校や進学校の親にも波及した、(2)進学校や進路多様校の親回収率の低下は、専門高校とは全く別のメカニズムによって生じている、のどちらかであろう。

#### 4 高校生データとの対応関係からみる親欠票群の特徴

仙台高校生調査では、親の学歴や職業といった属性はすべて親票で質問される。この方法だと、親の属性を高校生に質問する方法(たとえば、尾嶋編, 2001)に比べ、より正確な情報を得ることができる。ただし、親票が回収できないと情報は得られないため、親子間の関連の分析を行う場合、親欠票が生じたケースは基本的に分析から除外される。

データ分析において、このような情報のロスは望ましくない。しかし、この問題を逆に利用することで、親欠票群の特徴を探ること



注) 「構造的欠票」は親不在による欠票群, 「非構造的欠票」は親不在以外の欠票群。

図2 高校生の4年制大学進学希望率に対する父親学歴と欠票の影響

ができる。すなわち、親票が得られた高校生と、そうでない高校生の意識や行動を比較することで、親が欠票の高校生の特徴を分析し、その結果から、欠票群の親の特性を推測する、という方法である。

高校生の教育アスピレーションと親学歴の関係为例に、このことを検討しよう。分析に用いる変数の定義は以下の通りである。教育アスピレーションは、4年制大学進学希望の比率である。父親学歴は、有効票については「高卒」（高卒以下）と「大卒」（短大以上）の2カテゴリーとする。ここに父親欠票のカテゴリーを2つ追加する。1つは「構造的欠票」（父親不在による欠票）、もう1つは「非構造的欠票」（親不在以外の欠票）である。また、教育アスピレーションは高校種別の影響を強く受けるので、それをコントロールした上で分析を行う。高校種別は、前節で用いたものと同じである。

図2は、欠票群を含む父親学歴と高校生の教育アスピレーション（4年制大学進学希望率）

の関係を示したものである。

分析の結果は、以下のようにまとめることができる。

進学校では、有効票群と欠票群の間に大きな差がない時点（1994年、2003年、2007年）と、欠票群が比較的大きく落ち込んでいる時点（1986年、1999年）があり、一貫していない。しかし、全体的には欠票群の方が有効票群よりもアスピレーションが低めになる。

進路多様校では、各カテゴリーの間に比較的大きな差がある。有効票群で見ると、父親が大卒の高校生は、父親が高卒の場合より教育アスピレーションが高い。構造的欠票群のアスピレーションは、親が高卒の場合と同程度か、それよりも低い（2007年のみ高卒群を上回る）。非構造的欠票群の高校生のアスピレーションは、5時点中4時点で父親高卒群に近い値になる（1999年は大卒群に近い）。

専門高校は4年生大学進学希望率が全般的に低いので明確な差が把握しにくいだが、有効票群内のアスピレーションは、1999年と

2003年を除いて「父大卒>父高卒」となっている。一方、欠票群には一貫した傾向はみられない。

以上のように、父親欠票群の高校生の教育アスピレーションの関係はあまり明確ではないものの総じて低めであり、父親大卒の場合の教育アスピレーションを下回る場合が多い。ただし、構造的欠票群と非構造的欠票群では、その性質はやや異なると思われる。

まず構造的欠票群だが、早期（15歳時）の父親不在は子どもの教育達成を低めることが指摘されている（稲葉，2008；三輪，2005；三輪，2008）。父親不在の影響は、高校生が進学に伴う家計の経済的負担を考慮するといった形で、教育アスピレーションの形成段階ですでに存在するのだろう。

一方、非構造的欠票群は調査時点や高校種別によって挙動が異なるので、一般的な傾向を議論することは難しい。あえて解釈を施すとすれば、各グループの差が比較的明瞭な進路多様校については次のように考えることができるだろう。まず、父親高卒群の高校生より、父親大卒群の高校生の方が教育アスピレーションは高い。そして、非構造的欠票群の高校生の教育アスピレーションは、父親が高卒のグループに類似することが多い。したがって、非構造的欠票群の高校生は、父親高卒群の高校生に近い性質を有していると推測できる。高校生の教育アスピレーションに対する親学歴の効果については、「学歴の高い親ほど学歴の重要性を高く評価するため教育熱心であり、その結果として子どものアスピレーションも高くなる」と考えるのが妥当だろう。だとすれば、非構造的欠票群の父親は、(1)高卒の人が多く、(2)学歴とは関係なく子どもの教育に対する意識の面で高卒に近い父親が多い、のいずれか、あるいは両方である

と推測される。<sup>9</sup>このことは、3節2項で触れた高校種別による親の回収率の差異の問題ともリンクすると考えられる。

以上の考察が正しいとすると、親の有効回答のみを用いて分析した親子対応データの結果には偏りが生じる可能性がある。たとえば、非構造的欠票群の親が教育に熱心でないのだとすれば、親学歴と高校生の学習意識や学習行動の関係の分析を行った場合（たとえば、片瀬，2005；神林，2007；海野・片瀬編，2008など）、親の影響が過小に推定される可能性がある。したがって、仙台高校生調査のようなタイプの親子データを分析する場合は、親欠票群の影響に十分注意する必要がある。

## 5 むすび

学校調査において親欠票の生じる要因については、以下のようにまとめることができる。(1)親欠票は構造的要因と非構造的要因によって生じる、(2)構造的欠票、特に父親の構造的欠票は、離婚率の増大といった家族の変化によって、今後も増加すると予想される、(3)これに対して、非構造的欠票は、親の教育への関心が低いことによって生じている可能性がある、(4)構造的欠票・非構造的欠票とも、親子間の関連について分析をした場合、結果に偏りをもたらす可能性がある。

このうち構造的欠票の増加は、離婚率の増加という家族変動によるものであるから制御不能である。これに対して、非構造的欠票群が教育への関心の低さによって生じているのだとすれば、教育に対する関心が高まれば、回収率は上昇する可能性も考えられる。実際、図1によれば、2003年から07年にかけて仙台高校生調査の回収率は少なくとも低下傾向に歯止めがかかった。この時期は、いわゆる

「学力低下論争」(市川, 2002) が様々なメディアを通じて流通・消費された時期でもある。ここではその真偽には立ち入らないが、少なくともこうしたメディア情報は、進路選択時期を迎えた高校生の親にとって、教育への関心を高める契機になったと想像できる。

もしそうであるとするならば、学校調査における回収率を高めるうえでは、仙台高校生調査が行ってきたように、調査結果を学校や家庭にフィードバックし、調査対象者の教育に対する関心を喚起することで、回収率を回復させる途も拓かれてくる。山内らは、面接調査への非協力者に郵送調査を行い、非協力者のなかに「調査に対して疑問や抵抗感を持ち、調査に協力しなかったものの、何らかの条件次第では有効になる可能性があった人」(山内・米倉, 2002: 118) がいたことを指摘し、こうした「世論調査における『浮動層』」に働きかける必要性を強調している。また中田(2004)によると、学校においても、基本的に調査者を信頼し、調査の意義を認めているが、調査の方法・内容に関する適切な説明や事後報告(フィードバック)を欠いた場合、調査への疑問・不審・不満がもたれる、という。ただし、社会調査結果をフィードバックするにしても、調査の社会的意義は「その調査結果が社会的な関心を呼んでいることを認識することで実感できる」から、そのためには信頼できるメディアを通じて結果が報道されることが望まじらう(稲葉, 2006: 14)。佐藤(2008)も指摘するように、メディア(とくに低階層の情報弱者に愛好されるテレビ)が情報格差のセーフティネットになり、「一億総博知化」の潜在力をもつならば、こうしたメディアを通じて社会調査の結果とあわせて、その意義を伝え、ひいては調査の回収率や妥当性をあげるといふ再帰的なループを確立するこ

とこそ、今日の社会調査に求められていることではないだろうか。

注

- 1 なお、この調査データの分析結果については、(片瀬, 2005; 海野・片瀬編, 2008)などを参照のこと。
- 2 ただし、近年では学校も多忙になり、ホームルームや授業の時間で調査を行うことは難しくなっている。そのため、特に2007年調査では、高校生も実質的に配票調査になっている。
- 3 注2で述べたように、高校生票の回答は学校での集合式でなく(生徒が家庭に持ち帰っての)自記式で行われることが多くなっている。このことが高校生票の回収率低下の一因と考えられる。
- 4 家族構成の質問では、どういう場合に不在と判断するかは指示は特にないので、親が不在と回答されている場合、それが離死別によるものなのか、一時不在によるものなのかはデータからは識別できない。また、親が不在の場合は「親代わりの方」の回答を許容しているので、親が不在の場合でも有効な親票が存在するケースもある。表2以降の全ての分析では、「親不在となっているが親票が存在するケース」は分析から除外した。また、1986年調査と1994年調査の対象校には高専が含まれるが、全ての分析で高専は除外した。
- 5 高校生の親不在率 = 高校生票における「親はいない」という回答の比率。
- 6 たとえば、「人口動態統計」(厚生労働省)、「全国母子世帯等調査」(厚生労働省)。高校生の回収率が低下していることを考慮すると、実際の親不在率はさらに高いと思われる。
- 7 親欠票における親不在率 = 高校生票の「親はいない」という回答数 ÷ 親欠票数。なお、親欠票数 = 計画標本数 - 親有効票数、である。
- 8 なお、学校種別をコントロールした場合、それぞれの学校層における「親不在」の標本数はそれほど多くない(20前後となる場合が多い)。したがって、このカテゴリーの進学希望率は参考程度に解釈する必要があるだろう。
- 9 ただし、非構造的欠票群は親学歴と関係なく生じており、それゆえに高卒の親と大卒の親が混在する結果として父親高卒群と父親大卒群の中間的な位置になる、という可能性も否定できない。

文献

- 市川伸一, 2002, 『学力低下論争』筑摩書房(新書)。  
稲葉昭英, 2006, 『社会調査の困難』を考へる』『社会学年誌』47: 3-17。  
———, 2008, 『父のいない』子どもたちの教育達

- 成——父早期不在者・早期死別者のライフコース」  
中井美樹・杉野勇編『ライフコース・ライフスタイルから見た社会階層』（2005年SSM調査シリーズ9）2005年SSM調査研究会，1-19。
- 神林博史，2007，「高校生の学習時間とメディア接触時間——仙台圏の高校生データを用いた分析」『東北学院大学教養学部論集』147，1-22。
- 片瀬一男，2005，『夢の行方——高校生の教育・職業アスピレーションの変容』東北大学出版会。
- 三輪哲，2005，「父不在・無職層の帰結——将来の地位達成格差とその意味」尾嶋史章編『現代日本におけるジェンダーと社会階層に関する総合的研究』科学研究費補助金成果報告書，223-33。
- ，2008，「教育達成にみられる出身階層の影響」谷岡一郎・仁田道夫・岩井紀子編『日本人の意識と行動——日本版総合的社会調査 JGSS による分析』東京大学出版会，225-36。
- 中田敏夫，2004，「学校での調査の問題点——調査される側の論理」『日本語学』23(8)：146-54。
- 尾嶋史章編，2001，『現代高校生の計量社会学——進路・生活・世代』ミネルヴァ書房。
- 坂元慶行，2001，『日本の国民性調査』——社会調査研究のある最前線』『理論と方法』29：75-8。
- 佐藤卓己，2008，『テレビ的教養——一億総博知化への系譜』NTT出版。
- 海野道郎・片瀬一男編，2008，『〈失われた時代〉の高校生の意識』有斐閣。
- 山内利香・米倉律，2002，「調査不能の現状と課題——『あなたから見た世論調査②』から」『放送研究と調査』52(8)：110-21。





# 4

## 学校通しによる質問紙調査の可能性と 限界



### 1 はじめに

近年、教育界においてデータの重要性が高まったことを実感する。たとえば、国や自治体のレベルでは、教育政策や制度の立案において正当性を示すデータが欠かせなくなっているし、実施している教育行政の評価も数値で示すことが求められる<sup>1</sup>。学校現場では、学校経営の方針を検討したり指導改善のプランを立てたりするうえで必要となる資料を入手するために、児童・生徒の学力や学習状況が頻繁に調査されるようになった。さらに、多くの学校では、学校評価に用いるための調査が行われている<sup>2</sup>。

しかし、こうした学校をフィールドにした調査について、理論や方法をまとめた研究はそれほど多くない。1980年代にいくつかの研究が発表されているが、その後は学校をフィールドにするものの特殊性や課題が検討されることは少なかった<sup>3</sup>。調査そのものは一般的な社会調査の手法を援用して行われるので、それでも問題がなかったとも考えられる。とはいえ、実際に調査を行おうとする者（とくに初学者）にとっては、理論どおりにいかない部分について、葛藤を抱えながら試行錯誤をしてきたに違いない。これだけ学校をフィールドとした調査が普及している現状を考慮

木村 治生（ベネッセ教育研究開発センター教育調査室長）

すると、本来であればそこに生じる固有の問題を整理しておく必要があるだろう。

そこで、本稿では、学校をフィールドとして行う調査の可能性と限界を考察することにする。筆者自身はこれまで、おもに学校通しによる質問紙調査を手がけてきた。その対象は、児童・生徒、その保護者、教員、学校管理職などである。そのため今回は、こうした対象に対して学校通しの質問紙調査を実施する立場から、効果的な調査を実施するにはどうすればよいかという方法、実施の際にさまざまに生じる留意点などについて述べていきたい。

### 2 学校通しの調査の利点

教育をテーマにした調査を行う場合、学校をフィールドにする利点は大きいと考えられる。最初に、質問紙調査を学校通しで行うことの利点を検討する。それらは、①調査実施上の利点、②データ分析上の利点、③データ活用上の利点に大別できる。

#### ✿ 調査実施上の利点

調査を実施するうえでのもっとも大きな利点は、コストパフォーマンスの良さにある。学校をフィールドにした集合調査は、比較的少ない資源で、多くのサンプルを集めること



ができる。同じ自記式の質問紙調査でも、留置法や郵送法に比べると、はるかに小さい費用と労力で調査が可能だ。手軽さゆえに陥りやすい「畏」については後述するが、児童・生徒やその保護者を対象とした調査の場合、1校の協力が得られれば一度に100を超えるサンプルを集められる。

また、たとえば郵送法などでは、サンプルの抽出に使用する名簿が得にくくなったという環境変化もある。現在のところ住民基本台帳の閲覧は公益性の高い世論調査や学術調査には認められているが、個人情報取り扱いに関するルールは厳しくなることが予測される(長谷川, 2008)。教員や学校管理職を対象にした調査でも、以前は教職員名簿を用いるケースが多かったが、各地で目的外使用が問題になっており、安易な使用はできない。これらの点を考慮すると、教育にかかわる質問紙調査を行う場合、学校は必要なサンプルをもっとも効率的に得ることができる場であると言えるだろう。

### ❖ データ分析上の利点

次に、学校をフィールドにして得たデータは、分析するうえでどのような特徴をもつかを考えてみよう。学校通しの質問紙調査では、相当数の学校の協力を得なければ、特定の学校を選定したことによるバイアスが生じやすい。この点には十分な配慮が必要だ。しかし同時に、回収率が高いために、特定の属性をもつ者のみが協力することによって生じるバイアスを低減することができる。

郵送法では、どんなに回収率を上げる努力をしても3~5割の協力が得られれば良いほうである。こうした状況では、回答者に偏りが生じやすい。たとえば、保護者を対象とした調査では、仕事をもっていない教育熱心な

母親が多いと実感することがある。また、江上(2007)が指摘するように、留置法は調査員の技量に負うところが大きい。このため、調査を実施する際の手続きにおいてバイアスが発生する可能性が高くなる。さらに、調査員と対面することに抵抗感をもたない人が多くなるといった問題もあるだろう。

もちろん学校通しによる調査においても、無回答バイアスを絶無にすることはできない。とはいえ、児童・生徒を対象とした調査では欠席者を除くほぼ全員の回収を見込める。また、保護者を対象とした調査でも、経験的には7割程度の回収が可能だ。本当に困難な事情を抱えている保護者は脱落してしまうものの、回答者の偏りは是正される。教員調査の場合は、教育委員会や学校管理職の許諾を得てから行うケースと、抽出した学校に断りなく送付するケースで異なるが、前者の場合は高い回収率が期待できる。このように、データ分析の際に無回答バイアスを低減できることは、学校通しで調査を行う利点の一つである。

### ❖ データ活用上の利点

三つ目として、データ活用における利点を挙げておきたい。

調査そのものはさまざまな目的によって行われるが、学校をフィールドにすることから、事前に教員や学校管理職と打ち合わせをしたり、得られたデータを返却したりする作業が生じる。そうした一連の作業自体が、学校に対して実践的な意味をもつ。かりに調査が制度や政策といったレベルの関心によって行われたとしても、学校は得た資料から自校の特徴把握や課題発見につなげ、学校経営や子どもたちに対する指導の改善に役立てることができる。そうした実践への直接的な寄与は、

郵送法や留置法のように個別の対象者にアプローチする手法では難しい。もともとの調査目的を問わず、データ活用において学校への直接的な影響を与えることができるのは、学校通しの調査の大きな特徴と言えるだろう。

### 3 学校通しの調査の留意点

これまで述べてきたように、質問紙調査を学校通しで行うことにはいくつかの利点がある。しかし同時に、気をつけなければならないことも多い。手軽さゆえに陥りやすい「罨」を回避する工夫が求められる。以下では、留意点について考察する。

#### 🔍 調査実施上の留意点

1. サンプルングの問題 調査実施上の留意点として最初に考えなければならないのは、サンプルングについてである。学校をフィールドとした調査で悩むのは、「どの学校に協力をお願いするのが適切か」という問題だ。全国的な分布の実態や属性による分布の違いを調べる場合、サンプルング理論に基づいて一定の信頼性を確保しようと考え、相当数の学校に協力を依頼しなければならない。<sup>4</sup> 予算が潤沢で参加メンバーも多い大型プロジェクトであればよいが、実際には投下できる資源は限られるケースがほとんどだろう。学校通しの質問紙調査では、サンプルの代表性にこだわりすぎると実施が困難になる欠点がある。

それでは、代表性に欠けるから実施の意味がないかといえば、必ずしもそうではないだろう。盛山（2004）は、サンプルの選択について重要なのは、「代表性ではなく適切さだ」と述べている。この前提に立てば、探求目的にとって適切に関連する対象を選べばよいこ

とになる。実際に、学校通しの調査では、有意抽出で実践的な課題を検討するための調査が多く行われている。この場合、事後的にであっても仮説に沿った弁別（たとえば、特定の課題のある学校とない学校など）ができる対象であるかどうか重要になる。

さらに盛山は、「すべての変数とすべての個体データを得ることはできない」としたうえで、データに含まれていない属性の個体は分析できないものの、属性による違いは無視してもかまわないとリーズナブルに想定できれば、含まれていない属性にも妥当するだろうと想定できると指摘する。妥当と推察する根拠があれば、調査しなかった範囲にまでデータを汎用することが可能ということである。結果を安易に一般化することは控えなければならないが、学校通しの調査ではサンプルの代表性の確保に大きなコストがかかるだけに重要な指摘である。その際の妥当性を判断するためにも、どのような特性をもつサンプルかを強く意識して調査を行うことが必要だろう。また、実際には、さまざまな制限から理想的なサンプルングを行えないことが多い。探求目的にとって最善のサンプルを集める努力をしたうえで、予定どおりにいかなかった部分を含めてその過程や結果を開示する姿勢が求められる。

2. 学校に負担の少ない設計 調査を行うときに、なるべく学校に負担のない設計にすることも重要だ。近年、教員の業務量が増え、学校現場が多忙化しているという指摘がなされている。<sup>5</sup> そのようななかで協力を得ることから、学校をフィールドにした調査のノウハウや負担軽減の工夫が整理されてしかるべきだと思うが、初学者が参考になるものは残念ながら少ない。

負担を軽減する方法としては、大きく2つ

ある。1つは、協力する学校に実施の流れを明快に示すことである。調査票の到着から、調査の実施、回収までの流れが滞りなく行われるように設計し、学校管理職や担当教員の理解を得ておく必要がある。もう1つは、調査実施そのものの負担を軽減する配慮である。児童・生徒調査の場合は、なるべく短い時間で全員が完了できる分量とし、発達段階相応の表現レベルでなければならない。また、保護者調査の場合は、趣旨やデータの扱いなどに関する説明文書を添付し、問い合わせの電話窓口を設置する。調査について、学校が直接、保護者の質問に回答するのではなく、調査の実施者が対応するかたちにしたい。とにかく、負担軽減のための細かい配慮をしてはじめて調査が実現できる。

**3. 個人情報保護への配慮** 学校通しによる質問紙調査では、個人情報を取得せずに実施することが可能である。これは、利点の1つとして考えてよいかもしれない。調査は無記名を原則として、個人が特定できないかたちで実施するのが望ましい。学力調査など個人を特定しなければならないケースでも、個票を番号で管理するなど、実施者には個人が特定できないかたちで調査を行う方法を考える。また、対象者に対する配慮として、糊づけの封筒に入れて回収することをお勧めする。これは、調査の中間段階で実施者以外の第三者の目に触れることがないという姿勢を示すためである。個人情報保護法では、大学等の研究機関に所属する者が学術研究目的で個人情報を扱う場合は、「個人情報取扱事業者の義務」が免除されると規定されている。しかし、重要なのは、法的に問題があるかどうかということではなく、調査対象者に個人情報が守られているという安心感をもってもらうことである。

やむをえず調査の過程で個人情報を入手しなければならぬ場合は、それをどのように扱うかについて取得から廃棄までのフローを作成し、学校に明示する必要があるだろう。この際に難しいのは、児童・生徒を対象とする調査の場合、保護者に許諾をとったり、個人情報の扱いについて報告したりすべきかという問題である。原則は、そうした問題をクリアして調査に臨むべきだといえる。ただし、保護者が調査協力を拒否したときに、その子どもだけ調査から排除をするのが困難といった課題も生じる。現在は学校が教育的な観点から必要と判断したものは協力が得られることも多く、学校が調査を受けるメリットを大きくすることで、間接的に子どもたちに利益を還元するしかない。しかし、今後は個人情報を取得しないで済む手法の開発と共有、子どもの個人情報の扱いに関する倫理規定の作成などが必要となろう。

**4. 学校へのフィードバック** 学校をフィールドにした調査では、学校が協力することの意義を感じてもらえる調査内容でなければならない。また、その結果返却も欠かせない作業だと考える。当該校の数値を全体と比べられるようなかたちでまとめたり、結果に対して専門の立場からコメントしたりするなど、教育実践に生かせるかたちでデータをフィードバックするようにしたい。たんにフィールドを借りるだけでなく、学校と相互に利益を得る関係を成立させることが、調査の前提となる。さらに、報告書などの社会的な発信のまとめも報告すべきである。協力した教員や管理職は、自分たちの協力の結果がどのように生かされるのかを知りたいと考えている。そうした義務を誠実に果たす姿勢が、調査の実施者に求められる。

## ❖ データ分析上の留意点

1. サンプルの偏りに対する配慮 つづけて、分析する際の留意点について検討しよう。先に、サンプリングの問題について考察した際、学校通しの質問紙調査では代表性にこだわりすぎると実施が困難になる欠点があると述べた。大きな予算がついた調査でもないかぎり、理想的なサンプリングを行えることは少ない。また、一定数の無作為抽出をしたとしても、依頼した学校のすべてに協力を得られるとも限らない。こうしたことから、実施者は、協力を得た学校に偏りがある可能性をつねに意識する必要がある。そのために、学校の外的特性（地域特性や学校の規模など）と内的特性（教員や教育実践上の特徴など）の両方を、何らかのかたちで把握できるようにしておきたい。それらの情報は、得られたデータがどの範囲まで汎用可能なのかを類推する手がかりとなる。

くわえて、先行調査の結果も、そうした類推に有効だ。たとえば、全国学力・学習状況調査（文部科学省）のようにほぼ全数に対して実施された調査や信頼の置けるサンプリングを行っている調査とまったく同じ質問を含めておけば、数値の比較によって協力校の特性をつかむことができる。学校通しの調査では、母体のなかで協力校（および対象者）がどのような位置にあるのかを認識したうえでデータを分析することが重要であり、解釈をより広い範囲に適用する場合の妥当性の吟味にも、こうした情報が必須となる。

2. 調査票と集計表の開示 調査票と集計表（全体集計や基礎的なクロス集計）の開示はデータ分析の後に行うべきことであるが、第三者が調査の全体像をつかんだり解釈の適切さを判断したりするためにも、極力実行した

い。どのような統計分析も、最も簡単な統計量（度数分布や平均など）を基礎にしており、ここから多く情報を得ることができる（松原編、1985）。また、調査そのものやデータの解釈には誤りや作為性につきものである。そうしたことを隠蔽しようとするのではなく、開示するなかから誤りや作為を見つけようとする姿勢が大切である。藤村（2007）は、調査の公表において「自分たちがしたことを、自分に分かるように、また他者に分かるように物理的な形にすること」の重要性を指摘しているが、調査票と集計表の開示はその根幹を成すものといえよう。

## 4 仕組みとして考えておくべきこと

以上、学校通しの質問紙調査の利点と留意点を述べてきた。最後に、学校をフィールドにした調査に関して、仕組みとして考えておくべき課題を挙げたい。

現在、学校では多くの調査が行われている。「学校に負担の少ない設計」の項でも触れたが、こうした調査を含めた事務的な負担が大きく、教員が子どもと向き合う時間を奪っているという指摘がある。2007年に答申された中央教育審議会の『今後の教員給与のあり方について』では、教員や学校管理職の負担を減らすために、「国・都道府県・市町村等が行う調査の縮減・統合を図る」ことが提言されている。このように、行政においては調査を減らすことが意識されはじめているが、研究目的の調査も自制的でなければならない。

そのために、調査を実施する側には、無駄な調査がなくなる仕組みを整える努力が求められる。そのもっとも有効な手立ては、データアーカイブの充実だろう。欧米に比べると、わが国では圧倒的にデータを共同利用すると

いう発想が弱い（吉野，2002）。新しい調査には、実施者の側に多大な労力と費用がかかるうえ、対象者の負担も大きい。個票データを管理し、研究者の間で共有できる仕組みができれば、それらのコストの低減を図ることができる。1つの調査をいろいろな形で活用することは、調査に協力した学校にとっても、無駄な調査をせずに済んだ学校にとっても有益といえる。

さらに、既存データの利用は、初学者の教育・訓練にも役立つ。学校通しの調査では、教育実践に直接寄与することが求められる。既存データを用いれば先行研究が行っている工夫を学んだり、新たな調査を行うときの比較可能な項目を探索したりすることができるが、このような経験は調査レベルを向上させるうえで必要である。初学者に教育の機会を与え、総体として研究のレベルアップを図り、教育実践に役立つデータを還元する——こうしたことは、調査に協力する学校に対する研究者の責務であると考えている。公的な資金を用いて行った調査であれば、なおさらだろう。

研究の適切性や解釈の妥当性を検証するうえで、データの開示は有効である。先に、調査そのものやデータの解釈には誤りや作為性がつきものであると述べた。大切なのは、その検証可能性を閉ざさないことだ。研究の枠組みが適切かということや、解釈に錯誤はないかということは、分析の過程や集計表・クロス表などからもおおそ推察できるが、個票データがあればより明確になる。また、盛山（2004）の指摘のとおり、筆者も社会調査では「代表性ではなく適切さ」が重要だと考えるが、何が適切かということには基準のあいまいさが残る。とくに学校通しの調査では、抽出した学校に偏りが生じることが多い。そのために、データの開示がより重要な意味

をもつ。

## 5 おわりに

物事を考えるにあたっては、まず状況を正確にとらえる必要がある。状況把握に、数値は有効なツールだ。同じ物差しで過去からの変化や他との比較ができれば、問題の特徴を比較的容易にとらえることができる。解決しなければならぬ課題について、当事者間で問題を共有するのにも便利だし、目標設定もしやすい。また、教育実践の成果を示すうえでも、数値を用いれば説得力が増す。上手に使いこなせれば、とても役立つ武器になることは間違いない。以前から、教育界にはあまりに印象論的な議論が多く、その議論を実証的に裏づけるきちんとしたデータが少ないと言われてきた。「子どもの目が輝いているから優れた実践だ」と言われても、目の輝きの程度はなかなか比べられないし、再現性もない。優れている点を数値で客観化できれば（それは相当難しいことではあるが）、その実践を共有のものにする道も見えてくる。この意味で、実証的なデータをもとに物事を考えようとする姿勢そのものに、誤りはないように思う。こうした状況を踏まえると、学校通しの質問紙調査で明らかにできることはまだまだたくさんあるし、実践への活用という点でも大きな可能性を有していると考えている。

しかし、その行き過ぎも心配である。数値で測れるものは、全体のごく一部であるということも認識しておきたい。また、計測可能なものでも、調査にはさまざまな限界があり、完璧に測れるものではない。そのようななかで、データを過信したり、過大に評価したりすれば、別の問題が発生するように思えてならない。少なくとも、フィールドとなる学校

には大きな負担をかけることになる。こうした限界をクリアするためにも、本稿で述べてきたような調査方法の共有、水準を維持する仕組みの整備、ルール・倫理の確立などが必要であると考え。関係諸学会や社会調査協会には、ぜひそのリード役を務めてもらいたい。

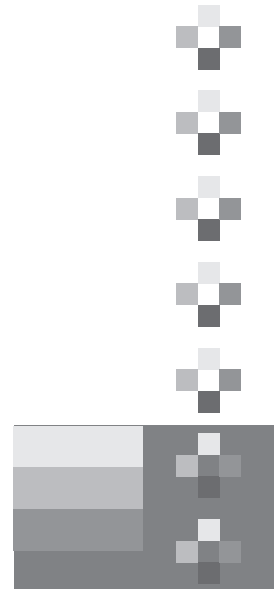
注

- ・1 たとえば、義務教育のあり方についての検討を集中的に行った中央教育審議会・義務教育特別部会の第1回の討議は、次のような委員の意見で始まる。「議論に入る前に、この審議会のスタンスを提案したい。この部会では、人とお金を巡った議論が行われるはずだが、その際は、エビデンスベース (Evidence base)、すなわち精緻な数値データをもとに議論しないと国民や他省庁等を説得できないので、可能な限りデータの提供をお願いしたい」。このように、近年では政策立案や他者への説明・説得にデータが必要だという認識が広まっている。
- ・2 文部科学省が実施した「学校評価及び情報提供の実施状況調査」によると、2006 (平成 18) 年度間に全公立学校 (幼稚園から高等学校まで) において学校評価のために外部アンケートを実施した比率は、88.9% である。
- ・3 松原治郎による編著 (1985) は実践的な性格が強く、学校をフィールドにすることの特殊性や課題が詳述されている。また、深谷 (1984-86) は3年間 (計 32 回) にわたって、学校をフィールドにした調査レポート『モノグラフ小学生ナウ』のなかで、「講座・子ども調査入門」と題する連載を行っている。その後、「教育調査」「学校調査」等のタイトルがついた著作は、管見の範囲では高島 (1998) のみである。
- ・4 たとえば、「平成 15 年度小・中学校教育課程実施状況調査」(国立教育政策研究所) では、小学校の 15% (3,554 校)、中学校の 23% (2,584 校) が抽出されている。また、「国際数学・理科教育動向調査 2003」(国立教育政策研究所) では、小学校・中学校をそれぞれ 150 校抽出することが、国際比較の条件になっている。
- ・5 「平成 18 年度文部科学省委託調査研究報告書・教育勤務実態調査 (小・中学校) 報告書」(国立大学法人東京大学, 2007), 「第 4 回学習指導基本調査報告書」(ベネッセ教育研究開発センター, 2008) など。
- ・6 東京大学社会科学研究所の SSJDA (Social Science Japan Data Archive) も、そのホームページ

のなかでデータアーカイブの意義として、①二次分析によって多くの新しい研究が可能となること、②多数の努力成果である個票データを有効活用により調査を受ける方の負担を減らすこと、③わが国の統計調査や社会調査の質の維持・向上に役立つことの 3 点を挙げている。(http://ssjda.iss.u-tokyo.ac.jp/about.html, 2008. 12. 10)

文献

- 江上渉, 2007, 「社会調査はどのように進めるか」森岡清志編『ガイドブック社会調査 (第 2 版)』日本評論社, 85-110。
- 藤村正之, 2007, 「公表の方法と報告書作成の要領」森岡清志編『ガイドブック社会調査 (第 2 版)』日本評論社, 293-321。
- 深谷昌志, 1984-86, 「講座・子ども調査入門」(32 回の連載)『モノグラフ小学生ナウ』Vol. 4-1~6-12。
- 長谷川公一, 2008, 「調査倫理と住民基本台帳閲覧問題」『社会と調査』1: 23-28。
- 松原治郎編, 1985, 『教育調査法』有斐閣。
- 盛山和夫, 2004, 『社会調査法入門』有斐閣。
- 高島秀樹, 1998 (改訂 2 版 2004), 『教育調査——教育の科学的認識をめざして』明星大学出版部。
- 吉野諒三, 2002, 「データオーガニゼーションの重要性とあり方」林知己夫編『社会調査ハンドブック』朝倉書店, 553-565。





## 社会調査としての学力テスト

——メイン NAEP に採用されている社会調査技法について——



村木 英治 (東北大学大学院教育情報学研究所・教育部教授)

全米学力調査 (NAEP: The National Assessment of Educational Progress) は、アメリカ教育省が教育テストサービス (ETS: Educational Testing Service) にその実施を委託して行っている。NAEP には、教育改革の効果やその調査時点において議論されている教育問題に焦点を合わせ全国的な学力傾向を調査することによって、その結果を行政担当者の教育政策の立案に役立てることを目的とするメイン NAEP (Main NAEP) 調査と、学力の長期的変化を追跡することを目的とするトレンド NAEP (Trend NAEP: Long Term Trend NAEP) 調査がある。

メイン NAEP では、英語読解、英語作文、数学、理科、合衆国史、公民、地理などの小・中・高校のカリキュラムに網羅されている教科科目がその調査対象となっている。近年では、音楽や演劇などのパフォーマンスアートもその対象に含まれるようになっていく。メイン NAEP は、これらの科目から毎年あるいは1年おきに1~3教科程度を選んで実施される。調査対象学年は4, 8, 12学年である。トレンド NAEP では、調査対象は9歳 (4学年)、13歳 (8学年)、17歳 (12学年) と年齢で層別されている。長期にわたっての学力変化の測定が目的のため、英語読解、英語作文、数学、理科の4教科について同じ問題を用い繰り返し4年周期で実施される。

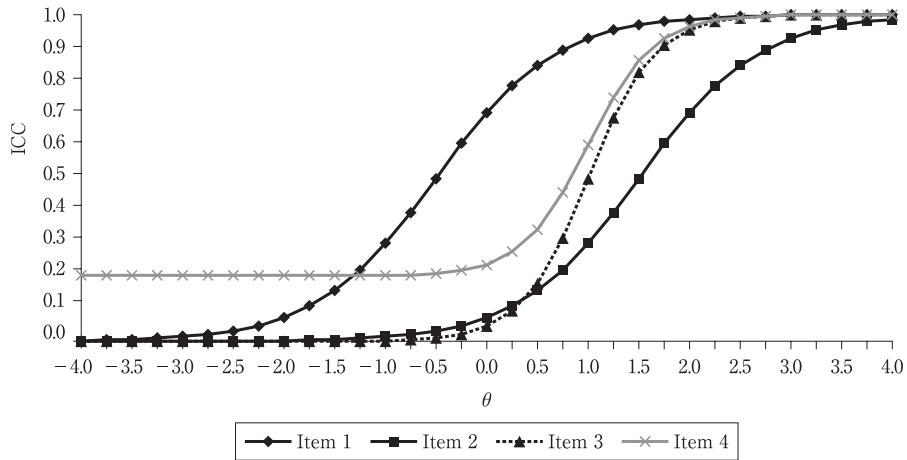
NAEP は、日本の「全国学力・学習状況調査」(2007年4月24日) が悉皆調査の形式をとるのと異なり、サンプル調査の姿勢を徹底している。被調査者である児童生徒を抽出するとともに、メイン NAEP では、テスト冊子に使われる項目自体もその背後に存在していると想定しているテスト項目の母集団からのサンプルであるとみなしている。それは、NAEP が1969年から始まった当初から、児童生徒の個々の学力を測定することにその目的はなく、あくまでも国勢調査のように、アメリカの人口を構成するさまざまな集団における学力の違いとその変遷を定期的に調べることに主眼をおいているという点にあるからである。社会調査としての NAEP の特徴として、前に述べたマトリックス標本調査法 (matrix sampling method) という形式があるが、それを可能としているのは項目応答理論 (IRT: item response theory) モデルの発展であり、また推算値法 (PV Technology: plausible value technology) の大規模なプロジェクトにおける利用である。ここではメイン NAEP に応用されているそれらの社会調査技法について順次紹介していく。

### 1

#### 項目応答理論モデル (Item Response Theory Model)

IRT モデルは、テストに含まれる各項目

図1 項目特性曲線：4例の項目  
Item Characteristic Curves



に対する正答（誤答）確率とそれに解答するのに必要と考えられる受験者の特性もしくは能力との関係を数理モデル（mathematical model）で表現したものである。受験者の特性や能力は直接観測することができない。それゆえ、それらを潜在特性（latent trait）とよび、通常 $\theta$ で表す。Birnbaum（1968）は次のロジスティック分布関数をとるIRTモデルを提案した。現在このIRTモデルが最もよく用いられている：

$$P_j(\theta) = c_j + (1 - c_j) \frac{\exp[a_j(\theta - b_j)]}{\exp[a_j(\theta - b_j)] + 1} \quad (1)$$

式(1)において、 $P_j(\theta)$ は潜在能力水準 $\theta$ でランダムに選ばれた受験者が項目 $j$ に正答する確率であり、このモデルは項目 $j$ を特徴づける3つ項目特性（item characteristics）パラメータ $a_j$ 、 $b_j$ 、 $c_j$ を含んでいるので、3パラメータロジスティック（3PL, three parameter logistic）IRTモデルとよばれる。パラメータ $b_j$ は項目の困難度（item difficulty）の指標であり、パラメータ $a_j$ は項目識別力（item discrimination power）、 $c_j$ は当て推量（guessing）パラメータとよばれる。IRTモデルはそれ

表1 4例の項目特性パラメータ値

	項目特性パラメータ		
	困難度 $b$	識別度 $a$	当て推量 $c$
項目1	-0.5	1.0	0.0
項目2	1.5	1.0	0.0
項目3	1.0	1.8	0.0
項目4	0.8	1.8	0.2

が含む項目パラメータの数で分類される。2PLモデルはパラメータ $a_j$ と $b_j$ を含み、最も単純な1PLモデルは $b_j$ のみを含む。1PLモデルはまたRaschモデルとも呼ばれる。

潜在特性 $\theta$ を独立変数として、式(1)の確率関数をグラフで表したものを項目特性曲線（item characteristic curve, ICC）と呼ぶ。図1に4つの項目のICCを例としてプロットした。それらの項目特性パラメータの値は表1に示した。

項目困難度 $b$ は $\theta$ 尺度上のICCの位置を表す。 $\theta$ 値が高くなるほど、その受験者の学力が高いとすると、項目2に正解するためには、項目1に比べてより高い学力が必要であるといえる。したがって項目2は項目1より難しい。学力が増せば、テスト項目の正解確率が増えていくから、項目識別度の閾値は正值（ $a > 0$ ）でありうる。パラメータ $a$ は、単



調増加関数 ICC の傾斜の度合であり、 $\theta = b_j$  の点における確率曲線  $P_j(\theta)$  の傾きの指標である。項目識別度  $a = 1.8$  をもつ項目 3 と  $a = 1.0$  の項目 2 を比べると、識別度が高ければ高いほど、同じ隔たりをもった 2 人の能力値の間の違いが、その正解確率の違いに敏感に反映する。言い換えれば、受験者の能力学力の違いをよりよく識別する。当て推量 (guessing) パラメータ  $c$  は、被験者が多肢選択型の問題において推量して答えたその解答が偶然正解である確率を表す ( $c \leq 0.5$ )。図 1 の項目 4 の ICC にみられるように、このパラメータは ICC の下方漸近線であり、正解確率における修正項である。この修正は、 $\theta$  が低い被験者により有効である。

IRT モデルのパラメータは特定の集団に依存しない項目統計量であり、またこのモデルを応用すると、テスト全体の難易度に依存しないで被験者の能力や学力を記述することができる利点がある。さまざまな集団、たとえば年齢別、性別、人種別、測定年度別などの集団ごとの結果を比較しやすいように、共通の尺度 (scale) を IRT は用意することができ、また少ない項目数でも十分に正確にその集団の学力 (proficiency) を推定できる方法を提供できる。また学力を総得点で表す場合、その閾値は限定されるが、学力パラメータ  $\theta$  がとりうる閾値は  $-\infty < \theta < \infty$  である。したがって、テストを等化 (equating) などをするることによって、あらゆる困難度をもつ項目群を同一  $\theta$  尺度上に理論上布置することができる。また同一尺度で表された項目パラメータを基に被験者の学力を算出すれば、同じテスト項目を回答しなくとも被験者の学力を比較することが可能となる。これらは IRT モデル応用の大きな利点である。

IRT モデルは古典的テスト理論のような、テスト全体と被験者集団を対象とした理論で

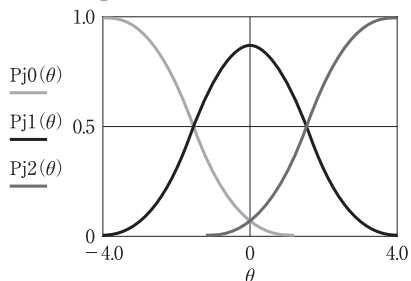
はなく、ひとりの被験者  $i$  がひとつのテスト項目  $j$  に対する反応をモデル化したものである。それゆえに、各被験者が答える項目が違ってくるとも知れないコンピュータ版テスト (computer-based testing, CBT) の実現には IRT 理論を欠くことができなかった。また後に述べるマトリックス標本調査法 (Matrix Sampling Method) の実施は、この IRT の応用なしでは困難である。

NAEP では一般化型部分採点モデル (GPC model: generalized partial credit model) という多値項目応答 (polytomous item response) IRT モデルも、テスト項目の分析に用いている。このモデルでは、学力水準  $\theta$  においてランダムに選ばれた被験者がテスト項目  $j$  で 0 から  $K$  までの順序尺度 (ordinal scale) のうちの  $k$  を選ぶ (あるいは  $k$  番目のカテゴリーをもってその項目の解答が評価される) 確率が

$$P_{jk}(\theta) = \frac{\exp\left[\sum_{v=1}^k a_j(\theta - b_{jk})\right]}{1 + \sum_{c=1}^K \exp\left[\sum_{v=1}^c a_j(\theta - b_{jk})\right]} \quad (2)$$

で与えられると仮定する。図 2 は、3 値反応の ICC の一例である。多値応答の場合は、パラメータ  $b_{jk}$  は項目ごとの反応カテゴリーの難易度を表すことになる。横軸に表された学力は、右にいくほど高くなる。さらに段階反応のカテゴリーが高くなるにつれて、より優れた反応と判断されている。学力が高い被

図 2 一般化型部分採点モデル (generalized partial credit model) の ICC



験者の項目応答ほど、高いカテゴリーに区分けされる。学力が低い場合、1番目のカテゴリー反応が出現する確率は、ほかのカテゴリーに比べ高い。そしてグラフでは学力が右に増加するに従って、つまり学力が増すに従い、1番目のカテゴリーの反応の確率が低くなり、2番目のカテゴリーの反応確率が高くなる。そして横軸のある一定の箇所まで、1番目の反応カテゴリーと2番目のカテゴリーに反応する、あるいはそのカテゴリーに評価される確率（odds）が五分五分だというわけである。さらに学力が高い被験者になると、3番目のカテゴリーに反応する確率が最初の2つのカテゴリーへの反応に比べて増加している。

NAEPでは、多肢選択式の項目には3パラメータロジスティックモデルが応用され、簡便型記述式項目の二値応答には2パラメータロジスティックモデルが使われている。作業型や拡張型記述式項目に対する反応は段階評価されるが、その際GPCモデルが分析に用いられる。項目応答理論（項目反応理論あるいは項目特性理論とも呼ばれる）の詳しい説明は池田（1994）、芝（1991）、豊田（2002）を参照してほしい。

## 2 マトリックス標本調査法 (Matrix Sampling Method)

被験者とテスト項目のどちらも標本抽出の原理に基づき選ぶデザインをマトリックス標本抽出法と呼んでいる。このマトリックス標本抽出法を採用することによって、被調査者である児童生徒が調査にさかねばならない時間を最小限に抑えながら、同時に広範囲な学力を測ることが可能となっている。

### ❖ 調査冊子作成の工夫：釣合型不完備ブロック計画

NAEP 学力調査のために作成された教科

表2 釣合型不完備ブロック計画の例 (1990年数学NAEP)

調査冊子のバージョン	配列1のブロック	配列2のブロック	配列3のブロック
1	A	B	D
2	B	C	E
3	C	D	F
4	D	E	G
5	E	F	A
6	F	G	B
7	G	A	C

科目ごとの問題設計に従ってテストを作成すると、問題冊子に含まなければならない項目数は相当なものになる。それだけの大規模な学力調査は、各被調査者には負担であるし、そのほかさまざまな困難が伴う。釣合型不完備ブロック計画（BIB design: balanced incomplete block design）に従ってそれらの問題を複数の調査冊子にわけける方法をNAEPはとっている。

NAEP測定時に児童生徒に配布される複数バージョンの学力調査冊子は、複数の項目セット（ブロック）から構成される。ブロック間では共通項目は設けない。BIBデザインでは、調査冊子間に「共通ブロック」を配置することによって、全体としての連結を保つ。どの冊子に含まれている項目も共通IRT尺度上においてモデルパラメータを推定できるために、この連結の確保は重大な条件となっている。表2には釣合型不完備ブロック計画の一例として1990年に使われたブロックデザインをあげた。

この数学調査のデザインでは、AからGまでの7つの項目ブロックを用い、その組み合わせ配列によって7種類の調査冊子が編集されている。たとえば調査冊子1と2ではブロックBが共通であり、調査冊子2と3ではブロックCが共通となっている。このようにしてどのバージョンもどれかのバージョンに連結しており、項目全体の母数が共通尺

度で表現されるようにデザインされている。ブロックデザインの詳細は各教科科目によって異なり、また新しい調査のたびに検討が加えられる。また表2でも明らかなようにブロックの小冊子の挿入順序がバージョンごとに異なる。このように問題項目の小冊子における出題順序をばらつかせることにより、不回答率の不均等を是正できる。さらにこれらの7種類の調査冊子を順に配布する（スパイラル）ことによって、特定の位置で受験する児童生徒間のテスト時に不正や、座席の位置による試験環境の条件などをコントロールすることもできる。

児童生徒に配布する調査冊子には教科科目のテスト項目とともに、児童生徒の各個人についての人種・民族、性別、家庭で使われている主な言語、保護者の最終学歴などの人口統計学的背景情報やその教科科目についての興味や勉強態度、家庭で読まれている活字媒体やTVあるいは宿題に費やす時間など学校内外での学習経験や学習環境についての質問項目が添付されている。被調査者への質問紙とともに公立か非公立かといった学校種、学校の位置する地域、学校のレベルや教育方針、さらには教師の学歴や教授資格など、校長や教師に回答対象となる質問紙も作成されている。これら被調査者、校長、教師への質問は、NAEPの調査結果の実際の教育政策にもたらす貢献度に大きく作用する。質問紙調査の質問作成には問題作成に劣らず深い配慮がなされている。

#### ❖ 被調査者抽出の工夫：多段階抽出法と層化抽出法

メインNAEP調査では被調査者抽出に層化抽出法（stratified sampling method）を用いている。この抽出法において、あらかじめ母集団をいくつかの部分集団に分けておく。こ

れら部分集団（枠 frame）を層（strata）という。これら層それぞれから独立に、必要な個数の標本を無作為に抽出する。この層化抽出法を使っている例としてアメリカの政治世論調査が良く知られている。政治世論調査はアメリカ合衆国の市民の多様性を反映するような調査結果を得るため、人種や宗教などさまざまな少数グループを含むようサンプルが抽出される必要がある。そのため人種別に層を形成したりして、それらの各層から必要な標本数を無作為に抽出する。こうすることによって、標本全体から単純無作為抽出するより、少数人種あるいは少数母集団からの標本を確実に得ることができる。より民意を反映している世論調査が実施されるというわけである。また各層を測定値においてできるだけ等質にすることによって、小さな標本で精度の高い母数を推定することもできる。

NAEPでは層化抽出法とともに多段階抽出法を用いて大きな枠から徐々に抽出枠を絞り込んでいく。まずメイン及びトレンドNAEPでは米国全土を北東地区、南東地区、中央地区、西部の4つの地域にほぼ同数の人口が含まれるように分け、各地区をさらに都市部と非都市部の2つに分ける。このように分けられた8つの集団を「ユニバース」とよぶ。これらのユニバースから第1次抽出単位がサンプルされる。ある地区では黒人やヒスパニック系の少数民族に属する生徒の在学の割合や各家庭の収入や学歴などの変数をもとにそれら抽出単位が決定される。このように、教育政策上重要な集団がその母集団における規模が小さいからといって標本にもれてしまわないように、NAEP調査の目的にかなった標本が最終的に抽出されるように層化抽出法を用いているのは政治世論調査と同様の考えからである。たとえば1996年度のメイン調査では94の、トレンド調査では52の第1

次抽出単位が抽出されている。

次のサンプリング段階ではこれら第1次に抽出された集団から第2次抽出単位としての学校が選出される。各第1次抽出単位内に存在する公立と非公立別の学校のリストを作成する。このリストをもとに、非公立学校や少数人種が多く在籍する学校を、母集団における実際の割合よりも多く抽出されるよう調節しながら、NAEP調査を依頼する学校を抽出する（過剰抽出）。このように選んだ学校のすべてからNAEP調査の承諾が得られるわけではない。これまでの記録では、20%ほどの学校が何らかの理由で不参加であったという。各年のNAEP調査の実態はNAEPテクニカルレポート（e.g. *The NAEP 1996 Technical Report*）に詳しい。

最終の抽出単位である児童生徒は各学校で作成された在籍簿から、少数人種やNAEP調査の目的に必要な少数集団に属する児童生徒が抽出にもれてしまわないよう集団ごとに異なる抽出確率を与えた上で無作為抽出をおこなう（過剰抽出）。このような抽出過程の結果として、各学年における母集団の0.2%ほどにあたる標本が抽出されるという。抽出単位の学校と同じく、NAEPでは選ばれた児童学生のNAEP調査参加について強制はできない。プロジェクトでは学校や個人の調査参加を確保するための努力を続けているが、不参加となった学校および個人の代わりとしてボランティアとしての参加は測定結果の精度を損なうため一切認めていない。

NAEPにおける少数集団の学力についての情報の確保は過剰抽出によって行われるが、母集団と標本集団における属性集団の過剰抽出によって生じた割合の違いは、データ分析の段階で「標本加重」によって修正される。このような標本抽出プロセスにおける工夫によって、集団全体ばかりでなく少数集団の学

力についても信頼性の高い情報が得られるのである。

### 3 推算値法 (Plausible Value Technology)

NAEPの目的は、児童生徒の個人的学力値を推定し比較することではない。その目的は全米児童生徒全体、あるいは属性集団の学力分布を推定することであり、またその集団学力値の経年変化を測定することである。このためNAEPではRubin (1987) に提唱された推算値法 (PV technology: plausible value technology) を用いてグループスコアを求めている。

IRTによる従来の能力値推定方法は、例えば個々の受験者に60項目かそれ以上の十分な項目数を与えて、最尤推定法などを用いて受験者の能力値 $\theta$ を推定することであった。この場合 $\theta$ の推定誤差 (uncertainty) は無視できるほどの大きさなので、受験者の能力推定値 $\hat{\theta}$ を真の能力値 $\theta$ とみなして、母集団や下位集団別の能力値分布を求めることができる。しかし、NAEPのような大規模な学力調査では、限られたテスト時間内に、広範囲な学習内容からテスト問題を出題する必要があり、一人ひとりの受験者に実施されるテスト項目数は、非常に少ないものになる。したがって、個人の能力値推定のための従来のIRTによる能力推定方法を、大規模学力調査における能力値推定に適応しようとする、次のような問題点が生じる。

第一に、個人の受験者の能力推定値に含まれる誤差は無視できない大きさとなり、この能力推定値を用いて推定される母集団の能力値分布は、真の能力値分布を反映したものとはならない可能性が大いにある。こうした問題点が生じることは、1984年のNAEP読解調査の初期の分析において実証

されている。

第二に、受験者ごとに項目数や形式、内容の異なるテスト冊子が実施される場合、測定誤差分布が異なり、母集団の能力値分布を正しく推定することができなくなる。古典的テスト理論を用いた伝統的標準テストにおいては、時間と集団を厳しくコントロールして平行検査を実施することによりこの問題を回避することができるが、現在のNAEPのテストデザインやデータ収集過程においてはこうした方法はとられていない。

受験者の負担を減らすために一人当たりの問題数を少なくし、幅広い学習内容から出題をするため受験者ごとに異なるテスト冊子を与えているBIBデザインにおいては、最尤推定法などを用いて受験者の能力値を点推定(point estimate)するIRTの能力値推定方法は、母集団の能力値分布を推定する方法としては適切な方法ではない。

こうした問題点を克服するために、NAEPでは、個人の能力値ではなく、母集団の能力値分布を正確に推定する方法として、PVという方法が考案された。PV法ではRubin(1987)の複数代入(multiple imputations)の方法が、潜在変数モデルに適用され、NAEPデータ分析の基礎となっている推算値が算出される。ここで、 $x_r$ を受験者 $r$ の学力テスト問題に対する解答ベクター、 $y_r$ をその受験者に該当する児童生徒、教員、校長へのアンケート調査項目に対する反応とおく。 $y_r$ は背景変数(background variables)ともいわれる。受験者 $r$ のPV  $i_m$ は、事後確率分布からランダム抽出することにより求めることができる。事後確率分布  $p(\theta_r|x_r, y_r, \Gamma, \Sigma)$  はベイズの定理を用いて、

$$\begin{aligned} p(\theta_r|x_r, y_r, \Gamma, \Sigma) &\propto P(x_r|\theta_r, y_r, \Gamma, \Sigma) \\ &\times p(\theta_r|y_r, \Gamma, \Sigma) \\ &= P(x_r|\theta_r) \times p(\theta_r|y_r, \Gamma, \Sigma) \end{aligned} \quad (3)$$

$P(x_r|\theta_r)$ は受験者 $r$ の真の能力値 $\theta_r$ が与えられたときの、この受験者の項目応答ベクター $x_r$ からIRTを基に導き出される尤度確率関数であり、 $x_r$ は受験者 $r$ の潜在特性 $\theta_r$ のみ依存し、背景変数 $y_r$ に対しては独立であるという仮定のもとで同じ式(3)の $P(x_r|\theta_r, y_r, \Gamma, \Sigma)$ に等しくなる。 $p(\theta_r|y_r, \Gamma, \Sigma)$ は質問紙調査項目の応答である背景変数 $y_r$ とパラメータ値 $\Gamma, \Sigma$ を条件として与えたときの、多変量同時密度確率関数である。メインNAEPの分析では、 $p(\theta|y, \Gamma, \Sigma)$ に対して正規分布が仮定される。 $\Sigma$ は分散共分散行列、 $\Gamma$ は回帰係数をもつ線形モデルによる平均であり、NAEPでは背景変数の完全なベクトルの数百回選択された主効果と2方向の交互作用から得られる約200の主成分から算出される。組み込まれた主成分は条件変数(conditional variables)であり、回帰係数 $y^c$ として表現される。次式(4)に表されるモデルは各科目のデータに適合している。

$$\theta = \Gamma' y^c + \varepsilon \quad (4)$$

$\varepsilon$ は平均0、分散共分散行列 $\Sigma$ を与えたとき、多変量正規分布をする。各受験者に用いられた背景変数の主成分は、(標準化後の)全背景変数の分散の90%を説明しているという。主成分分析(principal component analysis)を利用している理由は膨大な背景変数を直接回帰曲線に導入するとその係数が不安定になるためである。それを回避するために主成分分析と回帰分析とを併用した、この手法を主成分回帰分析(principal component analysis)とよぶ。この分析において、 $\Gamma$ は各行が1つの尺度の効果(effects)を含む行列であり、 $\Sigma$ は尺度間の残差の分散共分散(matrix variance covariance of residuals)である。

全米の児童生徒の能力値分布を推定するNAEPの分析過程において、PVは通常の意味での個人のテスト得点ではない。これまで

PVについて繰り返し述べてきているように、個人の受験者の能力値推定にとって最適な点推定 (point estimate) の累積から得られた分布は、母集団の特徴を推定するものとしては不適切であり、また不安定な推定値しか得ることができないという考え方は大変重要であると思われる。PV算出の具体的なアルゴリズムと用いられているコンピュータプログラムについては齊田・村木 (2005) に詳しい。

NAEPでは5つの推算値  $\hat{t}_m$  を式(3)の事後確率分布から抽出し、それらを用いて代表値  $t^*$  とその分散を算出する。それからの計算手順を示す。

1.  $M$  個の PV  $\hat{t}_m (m=1, 2, \dots, M)$  を抽出し、ジャックナイフ (jackknife) の分散推定を用いて、 $\hat{t}_m$  の標本分散を推定し  $U_m$  とする。
2.  $t$  の最終推定値を  $t^*$  とする。

$$t^* = \frac{\sum_{m=1}^M \hat{t}_m / M}{M} \quad (5)$$

3. 推算値の標本分散の平均を求め、標本誤差を推定する。

$$U^* = \frac{\sum_{m=1}^M U_m / M}{M} \quad (6)$$

4.  $\hat{t}_m$  の不偏分散を求め、代入 (imputation) 間の分散を推定する。

$$B = \frac{\sum_{m=1}^M (\hat{t}_m - t^*) / (M-1)}{M-1} \quad (7)$$

5. 最終の  $t^*$  の分散は次式で示されるように2つの部分の合計である。

$$V = U^* + (1+M^{-1})B \quad (8)$$

TIMSS (Trends in International Mathematics and Science Study) や PISA (Programme for International Student Assessment) のようなIRTを基礎においてデザインされた大規模学力テストはこれまで述べてきたNAEPの調査データ分析手法を踏襲している (Mazzeo, Lazer and Zieky, 2006)。これらの学

力測定 (assessment) は日本で行われている学力調査のデータ分析が児童生徒の個々の総合得点をデータ分析の基礎ユニットとしているのに対して、各調査者の得点の計算を経ず、PVという集団分布の代表値を求めることから集団得点学力調査 (group-score assessment) とよばれている。学力調査を一種の社会調査として捉えるなら、それら集団得点学力テストプロジェクトが採用しているマトリックス標本調査法 (Matrix Sampling Method) は広範囲な学力を短時間で調査できる点、とても効率よく、また被調査者をサンプルすることで調査の目的が最も経済的に達成できると思われる。そしてそのデザインの基で行う学力調査では、IRTとPVはごく自然な分析手法なのである。悉皆調査である日本の学力検査はいままで述べてきた手法のどれも採用していない。これらIRTベースの学力テスト調査の技法の応用を日本の学力調査にも強く勧めたい。

#### 文献

- Birnbaum, A, 1968, "Some Latent Trait Models and Their Use in Inferring an Examinee's Ability," F.M. Lord and M.R. Novick, Eds., *Statistical Theories of Mental Test Scores*, Reading, MA: Addison-Wesley, 397-479.
- 池田央, 1994, 『現代テスト理論』朝倉書店。
- Mazzeo, J., S.Lazer and M.J.Zieky, 2006, "Monitoring Educational Progress with Group-Score Assessments," R.Brennan, Ed., *Educational Measurement*, 4th ed., West Port, CT: American Council on Education and Praeger, 681-99.
- Rubin, D.B., 1987, *Multiple Imputation for Nonresponse in Surveys*, New York: John Wiley & Sons.
- 齊田智里・村木英治, 2005, 「調査結果のスケーリングと推算値」『全米学力調査 (NAEP) の研究』全米学力調査研究会, 61-70。
- 芝祐順, 1991, 『項目反応理論』東京大学出版会。
- 豊田秀樹, 2002, 『項目反応理論[入門編]』朝倉書店。