



RIETI Policy Discussion Paper Series 17-P-010

「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」 —エビデンスに基づく教育政策に向けて—

石川 善樹
株式会社 Habitech

伊藤 寛武
株式会社 Habitech

植村 理
慶応義塾大学

田端 紳
慶応義塾大学

外山 理沙子
慶応義塾大学

中室 牧子
慶応義塾大学

分寺 杏介
東京大学

星野 崇宏
慶応義塾大学

松岡 亮二
早稲田大学

山口 一大
東京大学



Research Institute of Economy, Trade & Industry, IAA

独立行政法人経済産業研究所
<http://www.rieti.go.jp/jp/>

「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」

—エビデンスに基づく教育政策に向けて—¹

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²

要 約

本稿では、就学期の子どもの認知能力を計測する目的で行われている「学力テスト」のあり方を展望し、今後の教育政策をエビデンスに基づくものに転換していくために、どのような学力テストのあり方が望ましいかを論じる。まずは国内外で実施された学力テストの動向を振り返り、その後平成 27 年から開始された「埼玉県・学力学習状況調査」を 1 つのモデルケースとして取り上げる。本稿の後半では、平成 27 年から開始された「埼玉県・学力学習状況調査」のデータの簡単な記述統計と分析を紹介するとともに、これと同時に行われた埼玉県戸田市での教員調査の結果とともに、今後の改善点についても論じる。

キーワード：ピア効果、埼玉県学力・学習状況調査、項目反応理論（IRT）

JEL Classification: I21, I26

RIETI ポリシー・ディスカッション・ペーパーは、RIETI の研究に関連して作成され、政策をめぐる議論にタイムリーに貢献することを目的としています。論文に述べられている見解は執筆者個人の責任で発表するものであり、所属する組織及び（独）経済産業研究所としての見解を示すものではありません。

¹ 本稿は、独立行政法人経済産業研究所におけるプロジェクト「医療・教育の質の計測とその決定要因に関する分析」の成果の一部である。また「埼玉県学力・学習状況調査のデータを活用した効果的な指導方法に関する調査研究」として埼玉県学力・学習状況調査データの分析研究業務委託を受けた。埼玉県教育委員会から多大な協力を得たことをここに感謝する。また、本稿の原案に対して、深尾京司一橋大学教授、乾友彦学習院大学教授ならびに経済産業研究所ディスカッション・ペーパー検討会の参加者から多くの有益なコメントを頂いた。ここに記して感謝の意を表したい

² Corresponding author: 慶應義塾大学総合政策学部 中室牧子 (makikon@sfc.keio.ac.jp)

1. 国内外で実施されてきた学力テストの現状

1-1. 集団の統計的特性を把握するための学力テスト

本稿の冒頭で、「学力テスト」とはどのようなものかについての定義を述べる。学力テストとはある個人の学力の状態を測定する検査のことである。学力テストには学校で行われているような定期考査、入学者や資格試験合格者の選抜を目的とするいわゆる「入試」や外部資格試験等も存在するが、ここでは国や自治体が教育政策の評価や学校資源の配分などに利用する目的で行われる学力テストに焦点をあてて議論する。

国や自治体が主導して行う学力テストには、学校で行われる定期考査や入試などの選抜試験とは異なる特徴が存在する。1つ目は、学力テストの対象者の抽出方法である。国や自治体を実施する学力テストは、個人への成績のフィードバックではなく、母集団の統計的特性を把握するという目的を持つ。これを達成するため、母集団の生徒全員を対象にした悉皆調査あるいは、個人や学校単位での標本調査が行われる³。2つ目は、テストの問題は、教科指

³ 「**悉皆調査**」は、全ての対象者の情報を入手でき、誤差のない正確な情報が手に入るという利点があるが、一方で実施のために多くの費用と手間がかかるという欠点がある。他方、「**標本調査**」では、対象となる母集団のうち一部の対象者について調査し、その結果から全体の傾向を推定する。この方法は手間や費用を抑えることができる利点がある反面、標本誤差が生じる可能性があるため、抽出する標本を偏りなく選ぶように注意しなければならない。ここで、偏りのない標本とは、統計的に母集団と性質が同じであると仮定できる一定数の標本を抽出することを指す（これを統計的サンプリングと呼ぶ）。学力テストの実施時には、「**層化多段階抽出法**」が用いられることが多い。層化多段階抽出法とは、母集団から無作為に標本を抽出するのではなく、あらかじめ設定された層（＝グループ）から、段階的に無作為抽出を行う手法である。例えば、まず全国の自治体の中から市町村を、さらにその中から学校を、さらにその中から生徒を無作為抽出するというのが層化多段階抽出法である。

導の専門家や心理測定学の専門家などの知見を動員し、生徒の学力をなるべく客観的かつ、異なる調査年や集団の間でも比較可能なように「測定」しているということである。この点については、1-3 で詳述する。

1-2. 教育政策の評価のために求められる情報

次にこうした学力テストの情報を用いて、どのような分析を行うことができるのかについて述べる。学力の決定要因は多岐にわたる。生徒本人の学習時間や学習意欲などはもちろん重要だが、そのみならず生徒やその保護者の社会経済的地位、教員の質、クラスメイトの属性、学校の設備や授業のカリキュラム、学校がある地域の環境などが学力に影響を与えていると見られる（Hanushek, 1989; Hanushek, 1997）。こうした状況をする際、経済学で標準的に用いられるモデルが「**教育生産関数**」（Education Production Function）である。これは、企業の投入物（資本、労働、技術など）と生産の関係をあらわしたもので、教育生産関数はこの生産関数の考え方を教育に応用したものである。つまり、学力を教育成果として教育への投入要素との関係を見たもので、この裏側には、生徒の学力を最大化させる投入要素の組み合わせとはどのようなものかを明らかにし、教育における最適な資源配分を達成しようという意図がある⁴。古くは、1960 年代に米国政府の主導で行われた「コールマンレポー

⁴ 過去、米国の公立小・中学校のデータを用いて、教育生産関数を用いて学校資源の効果について実証的な研究を行った研究を「開票法」（Vote Counting）によってまとめた Hanushek (1989; 1997)の研究では、学校資源（＝政府や州の学校への支出）は生徒の学力に殆ど影響を持たないことが示され、大き

ト」(Coleman et al 1966) に始まり、今日に至るまで教育生産関数を用いた研究は滂沱の数に上る。

一般的に、教育生産関数は次のように定義される。

$$A_{it} = f(X_i^{(t)}, SES_i^{(t)}, TQ_i^{(t)}, P_i^{(t)}, S_i^{(t)}, I_i)$$

Aは教育成果である学力である。左辺は全て投入要素をあらわしており、下表のように整理できる。添え字のiは個人を、tは時間をあらわす。

表 1：教育生産関数における代表的な投入要素

X (生徒本人の属性)	性別、学年、国籍、家庭での学習時間、学習意欲や勤勉さなど
SES (社会経済的地位)	保護者の収入、学歴、文化資本、社会関係資本、教育熱心さなど
TQ (教員の質)	教員の学歴、経験年数、非認知能力、指導法など
P (ピア効果)	クラスメイトや親しい友人の学力、属性、行動や習慣など
S (学校資源)	学校の設備やカリキュラム、生徒対教員比率など
I (生まれつきの能力)	時間を通じて変化しない個人の特性 (遺伝的な要因など)

このため、教育生産関数の推計のためには、教育成果である学力テストの結果のみならず、投入要素の情報も全て必要となる。別言すれば、学力テストの結果のみでは、「何が学力の決

な議論を呼び、一連の論争を「学校への支出は問題か」(“Money matters debate”)論争と呼ぶこともある (Hedges et al; 1994; Hanushek, 1994 など)。しかし、この後 Hanushek (1989; 1997) の用いた分析手法に対して、統計的な観点から様々な批判的検討がなされ (Greenwald et al, 1996; Heckman & Krueger, 2004 など)、最近の研究では、学校資源は学力に影響するということが示されている。しかし、Pritchett & Filmer (1998)が指摘するように、どのような学校資源であっても、必ず学力に影響するというものではなく、学校での資源配分が非効率に行われていれば、学校資源が学力に影響しないのはいうまでもない。このため、教育生産関数の推計を通じて、「どの学校資源が学力に影響しているのか」を把握することは、生徒の学力を最大化するような学校資源の配分を達成するという観点から極めて重要なのである。

定要因となっているのか」ということを明らかにすることができず、教育政策の評価や学校

資源の配分にとって有益な議論を展開することはできない。このため、上記の生産関数で投

入に該当するような情報は学力テストと同時に質問紙調査を行うなどして取得しておく必要がある。

とはいえ日本では、Oshio & Senoh (2007)がして指摘するように、学力データの研究目的での使用に制約があったことを主因として、教育生産関数の推定は最近になるまでほとんど行われてこなかった。北條 (2011) は後述する TIMSS のデータを用いて、教育生産関数を推定し、生徒個人の要因 (X) や保護者の社会経済的地位 (SES) などは学力の規定要因となっているものの、教員の質 (TQ) や学校資源 (S) などはほとんど影響を与えていないことを明らかにしている。学力の決定要因として、生徒個人や保護者の社会経済的地位の影響が大きいことは、日本の研究だけで示されているわけではなく、欧米の先行研究でも同様のことが示されている。当然これは重要な知見だが、学校教育や政策によって変えられない生徒個人や保護者の社会経済的地位の影響を取り除いた上で、学力に影響を与える要因を明らかにすることが重要である。例えば、前出の北條 (2011) は、習熟度別学級と学力の間に正の相関関係があることを明らかにしている。

特にこのときに注意すべきは、**保護者の社会経済的地位に関する情報**である。生徒に対して行う質問紙調査で、保護者の社会経済的地位について詳細な質問を行ったとしても、保護

者の状況についてよく承知していなかったり、記憶違いがあったりして、測定誤差が生じる可能性が高い。とはいえ、子どもの教育成果に保護者の社会経済的地位が影響しているという事は、一般に広く認知されるに至っていないため、政府や自治体が保護者に対して、子どもの調査に付随して、親の収入や学歴を尋ねる調査を行うことは極めてハードルが高いという現実がある。実際、過去の調査を見ても、子どもの学力調査に付随して、保護者に対する調査が行われたことは極めて稀である⁵。このため、保護者の社会経済的地位の代理変数として、家庭における文化資本をあらわす情報として「家庭における蔵書の数」などを尋ねている調査もある⁶。いずれにせよ、教育生産関数の重要な投入物である保護者の社会経済的地位についての充実した情報は、保護者の社会経済的地位と学力の関係を明らかにすると
いうこと以上に、保護者の社会経済的地位の影響を制御した上で、教員の質や学校資源が学力に与える効果を正確に推定するために必要なのである。

教育生産関数の推定にあたっては、重回帰分析の方法を用いるのが一般的だが、このとき、

⁵ 保護者の社会経済的地位が子どもの学力に与える影響について、わが国で行われたもっとも大規模な調査は、平成 25 年度に「全国学力・学習状況調査（きめ細かい調査）」中で、本体調査に加えて行われた「保護者に対する調査」である。これは「全国学力・学習状況調査」の対象となっている公立学校の中から小学校 430 校、中学校 414 校を対象に行われた調査であり、質問紙の有効回収率は小・中ともに 85%であった。この調査の分析でも、家庭の社会経済的地位と学力には相関関係があることが示されている（国立大学法人お茶の水大学, 2014）。

⁶ 例えば、Woessmann (2008)は TIMSS のデータを用いて、Ammermueller (2005)は PISA のデータを用いて、「家庭における本の数」と保護者の社会経済的地位をあらわす様々な変数と強い相関があることを明らかにしているほか、Kawaguchi (2016)は日本の PISA のデータを用いて同様のことを明らかにしている。

学力に影響を与えるような投入要素は、全て変数化されて推計に用いられていなければならない。もし、学力に影響を与えと考えられるにも関わらず、推計に用いられていない投入要素がある場合は、「脱落変数バイアス」(Omitted Variable Bias)を生じさせ、学力の決定要因を偏りなく明らかにすることはできない。このため、もしも質問紙の中に保護者の社会階層をあらわす変数やその代理変数が含まれていない場合は、教育生産関数を推計したとしても信頼性の高い分析結果はそもそも得られないということになる。行政判断にエビデンスを活かしていくためには、信頼性の高いエビデンスが蓄積されることが重要であり、そのためには、保護者の情報が必要である点について、国や自治体関係者および、調査対象となる保護者に対しても一層の理解を求めていく必要がある⁷。この点については、2—5でも改めて論じる。

またもう1つ重要なのが、調査票の有効回収率（＝調査票の回収数／調査対象者数）である。自治体が現状把握のために実施した質問紙調査などでは、生徒や保護者の自主的な回答を求めたものがほとんどであることから、回収率が非常に低くなっているケースも散見される。回収率が低くなってしまった場合は、統計的に母集団と性質が同じであるとは仮定できず、「外的妥当性」が低い（＝分析の結果が、この調査票に回答した人たち以外には当ては

⁷ こうした調査では、回答者の氏名や住所などの個人情報 は全て削除し、個人を識別する ID が割り当てられて分析が行われるなど、個人が特定されることがないように、最大限の注意が払われる。

まらず、一般化できない)と判断され、政策判断に用いるには不十分である。これ以外にも、特定の調査項目に回答拒否や無回答などが発生している場合は、欠測データとして扱われるが、この欠測データをどのように扱うかも慎重に判断する必要がある⁸。

1-3. 学力の「測定」を行う技術の発展

1-1 で、国や自治体が実施する学力テストの特徴として、学力を「測定」している点に触れた。近年、その「測定」の方法は、技術的に飛躍的な発展を遂げている。その1つとして、「項目反応理論」(Item Response Theory, 以下 IRT)がある(詳細は、Embretson & Reise, 2000)。これはこれまでに利用されていた「古典的テスト理論」(Classical Test Theory: 以下 CTT)の欠点を克服した新しいテスト理論として、TOEFL など入試や選抜などの重要な意思決定にかかわる(いわゆるハイステークスな)テストに利用されている。日本においても、多くの書籍が刊行され(e.g., 村木, 2011; 加藤・山田・川端, 2014)、国内の公的な統一試験としても情報処理技術者試験や医療系大学間共用試験などその利用は年々広がってきている。

IRT モデルが提案されるまでに一般的に利用されていた CTT ではテスト項目得点の和や正答数を測定の対象となる真値(能力値や特性値)と測定誤差の和によって表現しており、各項目は同じ重みで考えている。具体的には素点(x)と真値(t)と測定誤差(e)の3項を用い

⁸ 欠測データの取り扱いについては、星野・岡田(2016)を参照。

て、

$$x = t + e \quad (1)$$

と表現し、誤差 e は真値 t や素点 x とは無相関であると仮定する。CTT では真値の代理変数として素点を利用するという点では日常的に行われているテスト得点の利用と何ら変わらない。但し CTT ではその前段階として、例えば素点の分散に占める真値の分散を表す信頼性係数の評価を行い、信頼性が低い場合には利用する項目を変える、などというように素点を代理変数として利用することの正当性を担保するための様々な方法論が開発されている。

CTT の問題点は大きく 3 点あげることができる。まず(1)項目によって難易度や特性値・能力に対する識別能力に差があるということを見逃してきた。また(2)特性・能力の真値と項目の難易度や識別能力を分離できないために、たとえば得点が低いという場合に受験者の能力特性値が低いのか、難易度が高い（または極端な項目である）のかがわからない。その結果、(3)別の受験者・回答者集団間での特性値・能力値の比較ができない。

学校現場や受験産業ではしばしば素点を標準化したり偏差値に変換したりすることで集団間の比較が行われることがあるが、テストの難易度と集団間の能力分布の差が分離できないという点で CTT と問題点を共有している。

これに対して、IRT では「被験者の特性値・能力母数」と、(困難度や識別力といった)「項目の特性に関連する母数（項目母数）」が別個に定義されている。これにより、項目ご

との難易度や識別能力の違いが明確化でき、また被験者の特性値・能力母数と項目に関連する母数を別々に推定することが可能になる。

以下、具体的に正答／誤答型の 2 値の項目反応に対する IRT モデルである 3 パラメータロジスティックモデルを説明する。ある個人*i*の項目*j*に正答する確率 $\Pr(x_{ij} = 1)$ を、高ければ高いほど項目への正答確率が上昇する個人の能力値を表すパラメータ θ_i と項目を表すパラメータ a_j, b_j, c_j によって

$$\Pr(x_{ij} = 1 | \theta_i, a_j, b_j, c_j) = c_j + (1 - c_j) \times \frac{1}{1 + \exp(-a_j(\theta_i - b_j))} \quad (2)$$

と表す。 a_j は識別力パラメータとよばれ、ある特定の能力値前後の能力を区別できる度合いの強さをあらわす。 b_j は困難度を表すパラメータで項目の難しさを反映している。 c_j は当て推量パラメータと呼ばれ、多肢選択式のテストなどでのあて勘による正答確率を表すことに利用される。埼玉県の詳細解析では $c_j = 0$ とし、 a_j を項目間で一定と仮定した最も単純な 1 パラメータロジスティックモデルが利用されている。

IRT を採用することの大きなメリットとして、特定の項目を共通にする（共通項目デザイン）あるいは受験者・回答者の一部を共通にする（共通被験者デザイン）ことで、異なる時点で行われたテスト間の比較、複数の受験者・回答者集団間や異なる年齢間での特性値・能力値の比較といったことが可能になる。これをテストの等化という。

等化とは、複数のテストの目盛りを合わせ相互の結果を比較可能にする操作である。これ

により、TOEFL などでは(難易度や受験者集団の)異なる試験機会ごとの結果を比較可能にすることが可能になる。しかし、IRT モデルの利用には十分な準備や仮定が必要となる。まず、IRT を利用するためには、測定している能力が 1 次元である必要があり、つまりテストが 1 つの能力のみを反映していることが望まれるのである。また、等化など IRT の利点を利用するためには問題項目は基本的には秘匿されていなければならない。項目の難易度などがテストの項目を公開してしまうことによって変わってしまうことを防ぐことが必要だからである。また、IRT モデルのパラメータの推定にはそれなりに大きなサンプルサイズが必要である。このように、IRT モデルの利用にはいくつかの制限があることも留意する必要があるだろう。

1-4. パネルデータとクロスセクションデータ

いずれにせよ、こうした学力テストの意義は、単年度の結果のみに注目せず、継続して長期的な傾向を把握することにある。このときに、ある特定の学年を対象に、年度を越えて把握するような調査を「クロスセクションデータ」と呼ぶ。一方、同一対象の成長を追いかけて調査するものを「パネルデータ」または「追跡調査」などと呼ぶ。近年は、後者のパネルデータの有用性が強調されることが多い⁹。「教育の効果はすぐは現れない」という実感は、

⁹ 詳細は、北村（2006）、北村（2013）を参照。北村（2013）は、最近の米国の一流雑誌に掲載された労働経済学分野の論文を概観し、使用されたデータがクロスセクションデータである論文は概ね横ばいである一方、パネルデータである論文が急速に増加していることなどを例に挙げている。

教育の需要者側、供給者側双方にあらうが、クロスセクションデータではそのような「時間を通じた変化」を推し量ることができない。もし仮に、教育政策の効果が短期的には観察できなかつたとしても、長期的には観察できるような性質をもつのであれば、この「時間を通じた変化」というのは、教育政策の評価を行う上で極めて重要な情報だと考えられる。

また、やや専門的な用語を使えば、教育に影響を与えると考えられる時間を通じて不変の「観察不可能な異質性」(Unobserved Heterogeneity)を制御することができる点も重要である。例えば、教育生産関数の重要な投入要素の1つである I (時間を通じて変化しない個人の特性) はたとえ、質問紙調査を行ったとしても、適切に数値化することは難しい。これ以外にも、保護者の教育熱心さのように、研究者からみると観察することができないものの、教育成果に大きな影響を与えると考えられる投入要素もある。しかし、パネルデータを用いれば、こうした観察不可能な異質性の影響を制御して、教育生産関数を推計することが可能になる。

パネルデータは、同一個人を追跡したものだけとは限らない。学校や自治体を単位として、経年で調査を行えばそれもまたパネルデータであると考えられる。しかし、個人を学校や自治体を単位としたデータを分析する際には、「集計」がもたらすバイアスにも注意が必要だ。俗に言う「エコロジカル・ファラシー」(Ecological Fallacy)の問題である。特に、Hanushek et al (1996)は、個人を単位として推計した教育生産関数と、学校や自治体を単位として推計

した教育生産関数では結果が異なる場合があります、学校や自治体を単位とした教育生産関数は、学校資源が学力に与える効果を過大に推定している可能性があることを指摘している。このため、大規模かつ長期の追跡調査には費用がかかることが多いが、できる限り個人の単位で分析することが望ましい。

しかし、パネルデータの実施には困難が伴う。同一個人を追跡する途中で、引越しや病気などによって、同一個人から回答を得られなくなることも多い（このように個人の回答を得られなくなることによってサンプルが調査から脱落してしまうと、パネルデータは「不完備になる」という）。

現在、日本で政府が実施している未成年の子どもを対象にしたパネルデータの中でもっとも歴史が長いのは、厚生労働省の「21 世紀出生児縦断調査」であるが、この調査の場合 2001 年に開始され、現在も同一個人を継続して追跡が行われている。この調査は 2001 年には 47,015 人から回答を得ていたが、直近の 2015 年には 29,506 人からしか得られておらず、14 年が経過してなお 2001 年回答者の 63%を捕捉することができている。これは政府が実施しているからこそ追跡率の高さでもあり、これと同様のことを研究機関が独自予算で実施すると同調査と同程度の捕捉率を維持することは極めて困難であると考えられる。この意味において、大規模かつ長期の追跡調査は、インフラとして政府が手がけるべき事業であるといえよう。

前に、IRT では、複数時点のテストの等化を行うことを説明したが、この等化（これを垂直等化と呼ぶ）を行うことによって、個人の能力の経年比較も行うことができる。これによって、個人の能力の「時間を通じた変化」を観察することができ、さらに IRT モデルを用いて能力値を 1 次元上の尺度にすることにより、その変化分にどのような要因が寄与するのかを検証することが出来る。つまり教育生産関数における投入が、学力にどのような効果があったかという検証が可能になるのである。実は、学力をパネルデータとするためには、能力の経年比較が正確にできているということが保障されていなければならない、IRT はパネルデータとして設計された調査の中で用いられることによってその有用性を発揮する。

1-5. 日本で行われてきた学力テスト

ただし、日本では、IRT のような測定の新しい技術が取り入れられた政府や自治体主導による大規模な学力テストというのは、これまで行われてこなかった。本稿の後半で紹介をする「埼玉県学力・学習状況調査」はこの IRT を用いて、埼玉県下（ただし、さいたま市を除く）全ての公立小・中学校で行われた大規模な学力テストである。この学力テストについて紹介する前に、これまで日本で行われてきた学力テストの中で代表的なものとして、①「全国学力・学習状況調査」（文部科学省・国立教育政策研究所）や各自治体が独自に実施している学力テスト、②PISA や TIMSS などのような国際機関が実施している国際比較可能な学力調査、③慶應義塾大学パネルデータ設計・解析センターによる「日本子どもパネル調査」

を紹介したい。

1－5－1.「全国学力・学習状況調査」および自治体が独自に実施する学力テスト

2007 年以降、文部科学省が実施している日本全国の小・中学校の最高学年（小学 6 年生、中学 3 年生）を対象として行われている調査テストは「全国学力・学習状況調査」と呼ばれている。算数・数学と国語の 2 科目で、それぞれ知識力を問う問題(A)と知識活用力を問う問題(B)の 2 種類に分かれている¹⁰。また、学力テストだけでなく、児童・生徒の学習・生活環境の質問紙調査も同時に行われている。この学力テストは、悉皆調査で行われた年と標本調査で行われた年が混在するクロスセクションデータである。また、この学力テストは、「古典的テスト理論」に基づいて設計されており、テストの正答率がそのまま分析に用いられている。文部科学省は「平成 23 年度以降の全国的な学力調査の在り方に関する検討のまとめ」¹¹を公表し、悉皆調査で行うことや、調査結果については都道府県単位での集計データの公表に留める一方、個人成績や学校ごとの成績公表を行うか否かは市区町村教育委員会に委ねるという方針を定めている。この全国学力・学習状況調査の個票データに関しては、その実施の費用にかかる金額の大きさに鑑み（悉皆の場合約 60 億円／年、標本の場合 30 億円／年）、政府の内部のみで分析するのにとどまらず、研究者に研究目的での利用を求める声が

¹⁰ 2012 年(平成 24 年度)調査のみ理科も実施された。

¹¹ 「平成 23 年度以降の全国的な学力調査の在り方に関する検討まとめ」

(http://www.mext.go.jp/b_menu/shingi/chousa/shotou/074/toushin/1304351.htm)

あり、2016 年 9 月には文部科学省から「全国学力・学習状況調査個票データの貸与の在り方について（案）」¹²が出され、2017 年以降、研究者に研究目的で個票データの貸与が行われる方向で議論が進んでいる。

こうした国レベルでの学力テストに加えて、全国の都道府県教育委員会、指定都市教育委員会数のうち約 7 割が独自の学力テストを実施している¹³。その目的は、「児童・生徒一人一人の学習状況の改善に役立てるとともに、各学校における授業改善及び区市町村における学力向上のための施策の充実に役立てる」（東京都、2016¹⁴）などとされており、自治体の政策評価や教員の指導改善に活用している。文部科学省の「全国学力・学習状況調査」が独自に問題を開発しているのに対し、予算や技術的な制約もあり、自治体は独自問題を開発しているのではなく、大手事業者が販売している学力テストを購入し、実施しているところがほとんどと見られる。こうした市販の学力テストは、タイムリーに結果を分析し、政策や指導に活かすことができるというメリットはあるものの、他の自治体との比較や経年変化を正確に把握するのは難しい。また、学力テストに付随して、十分な質問紙調査を実施していないケースも散見されたり、行政担当者が大規模な個票データの分析に必要なスキルを身に付けて

¹² 「全国学力・学習状況調査個票データの貸与の在り方について（案）」

(http://www.mext.go.jp/b_menu/shingi/chousa/shotou/112/shiryo/attach/1377298.htm)

¹³ 平成 25 年度実施の都道府県・指定都市による独自の学力調査について

(http://www.mext.go.jp/a_menu/shotou/gakuryoku-chousa/sonota/detail/1335150.htm)

¹⁴ 東京都 H28 報道発表資料 (<http://www.kyoiku.metro.tokyo.jp/buka/shidou/chosa/pr160707.html>)

いないなどの理由から、学力の決定要因についての妥当な分析が行われているケースはほとんど見られない。

1－5－2. 国際的に比較可能な学力テスト（PISA／TIMSS）

国レベルでの教育水準等を調査・比較するためのテストの例として、経済協力開発機構（OECD）による「**国際的な生徒の学習到達度調査**」（Programme for International Student Assessment, 以下 PISA）¹⁵と国際教育到達度評価学会（IEA）による「**国際数学・理科教育動向調査**」（Trends in International Mathematics and Science Study, 以下 TIMSS）¹⁶がある。

これらの調査は、ともに （層化抽出法によって）抽出された標本調査のクロスセクションデータで、PISA で約 6,600 人、TIMSS で約 9,100 人（小 4：4400 人、中 2：4700 人）の生徒が対象となっている。PISA は、科学、読解、数学的リテラシーの 3 科目で、OECD 加盟国・非加盟国を含めて直近の 2015 年には 72 カ国が参加している。義務教育の終了段階にある 15 歳の生徒を対象にしており、2000 年以降 3 年ごとに行われている。生徒質問紙、教師質問紙、学校質問紙によるアンケートも付随して実施されている。TIMSS は、算数・数学と理科という理数系科目で、小学校 4 年生と、中学校 2 年生を対象にしており、1995 年以降 4 年毎に行われている。生徒質問紙、学校質問紙も実施されている。TIMSS は、前

¹⁵ 第 1 回調査は 2000 年、以後 3 年毎に調査している。最新の調査は 2015 年に行われた。

¹⁶ 第 1 回調査は 1995 年、以後 4 年毎に調査している。最新の調査は 2015 年に行われた。

回調査の小学 4 年生が 4 年後、中学 2 年生時になるという設計になっており、集団のレベルでは、「時間を通じた変化」を観察することができるようになっている。ただし、小学 4 年生時に調査を受ける児童と中学 2 年生時に調査を受ける対象は別であり、同一個人についてその変化を知ることができるわけではない点には留意が必要である。PISA と TIMSS は、学力の測定にあたって、ベイズ統計の枠組みを用いた「**推算値**」(Plausible Values, PV) と呼ばれる統計量を求め、集団の能力分布の分散を正確に推定することを試みている¹⁷。そして、PISA や TIMSS の個票データは、研究者に限らず誰でもアクセスが可能で、OECD や IEA のウェブサイトダウンロードすることができる。

1-5-3. 日本子どもパネル調査

政府や自治体、あるいは国際機関などの公的機関ではなく、大学が実施している学力テストが、慶應義塾大学パネルデータ設計・解析センターが実施している「日本子どもパネル調査」(Japan Child Panel Survey, 以下 JCPS)である。海外では、ミシガン大学に設置された Institute of Social Research (ISR)が、Panel Study of Income Dynamics (PSID)を実施するなど、大学がパネルデータの収集と公開に大きな役割を担っている。学力の測定に先駆的な方法が用いられていたり、分析を念頭において質問紙が設計されていたりすると、政策評価と

¹⁷ Plausible Value は測定誤差の影響を軽減し、集団としての能力分布を正確に推定するために利用される。PV も IRT の枠組みの技術である。詳細は、von Davier, Gonzalez, & Mislevy (2009)。

しても、研究としても精度の高いものになりやすい。日本でも慶應義塾大学パネルデータ設計・解析センター以外に、東京大学社会科学研究所附属社会調査・データアーカイブ研究センターや公益財団法人・家計経済研究所など複数の大学・研究機関が、大学や企業のリソースを活用して、パネルデータの収集にあたっているが、アメリカのように政府や自治体と大学が共同して、データの収集や公開などを行っている例は見られていない。

JCPS は全国から（層化抽出法によって）抽出された標本調査のパネルデータで、小学 1 年生から中学 3 年生までの約 1,000 人の児童・生徒を対象にしている（詳細は赤林・直井・敷島、2016 を参照）。2010 年に開始にされて以降、この調査の特徴として、学力テストに加えて、子どもを対象とした質問紙調査、保護者を対象とした質問紙調査も行われており、わが国で行われている学力テストの追跡調査としては画期的な設計となっている。一方で、利用可能な調査の中で有効回収率をみると、概ね 50～55%程度となっており、政府が実施している「21 世紀出生児縦断調査」などのパネルデータに比べると低くなっている（同調査の回収率は過去 14 年間、平均して 90%前後で推移している）。一方、このパネルデータは、非営利・学術目的での利用に限り研究者に無償でデータの提供を行っている。

次章では、こうした学力テストの動向を踏まえ、埼玉県で 2015 年度から新たに実施されている「埼玉県学力・学習状況調査」について詳細を述べる（埼玉県学力・学習状況調査と代表的な学力テストの比較は表 2 を参照）。

表 2：過去にわが国で実施された代表的な学力テスト

	全国学調	PISA／TIMSS	JCPS	埼玉県学調
測定の方法	CTT	IRT	CTT	IRT
調査方法	悉皆	抽出	抽出	悉皆
対象学年	小 6、中 3	PISA: 15 歳 TIMSS：小 4、中 2	小 1 ～中 3	小 4～中 3
対象者数	約 224 万人	PISA：6,600 人 TIMSS：9,100 人	約 1,000 人	約 30 万人
データ	クロスセクション	クロスセクション	パネル	パネル
公開	非公開	公開	研究者に公開	研究者に公開

（注）対象者数は全て直近の調査。

2. 埼玉県学力・学習状況調査の概要

2-1. 埼玉県学力・学習状況調査の対象と実施方法

埼玉県では、2015 年および 2016 年に、埼玉県下（ただし、さいたま市を除く）の全ての公立小・中学校の小学 4 年生から中学 3 年生の生徒に対して、埼玉県教育委員会が実施した「埼玉県学力・学習状況調査」を実施している。

表 3：埼玉県学調の参加自治体・学校数

	自治体数	小学校数	生徒数	中学校数	生徒数
2015 年度	62	708	150,310	360	148,013
2016 年度	62	708	149,227	356	146,323

（出所）埼玉県（2016）

埼玉県学調では、国語、算数・数学、英語といった 3 教科学力テストの他に、児童の非認知能力、学習への意欲や態度、家庭での学習状況、保護者の社会経済的地位などを尋ねた質問紙調査が、新年度が始まる 4 月初旬に各学校で教職員の監督の下、一斉に行われた。また、学校や市町村教育委員会での施策を尋ねた質問紙調査（オンラインで回答）も行われている。このため、後述するように、テストが実施された当日の欠席者（2016 年調査で平均 2～3%）を除いては全員が参加している。

この調査は、2015 年度に開始してからまだ 2 年しか経過しないものの、悉皆調査のパネルデータであり、生徒 1 人 1 人の学力が IRT を用いて推定されている。このため、古典的な

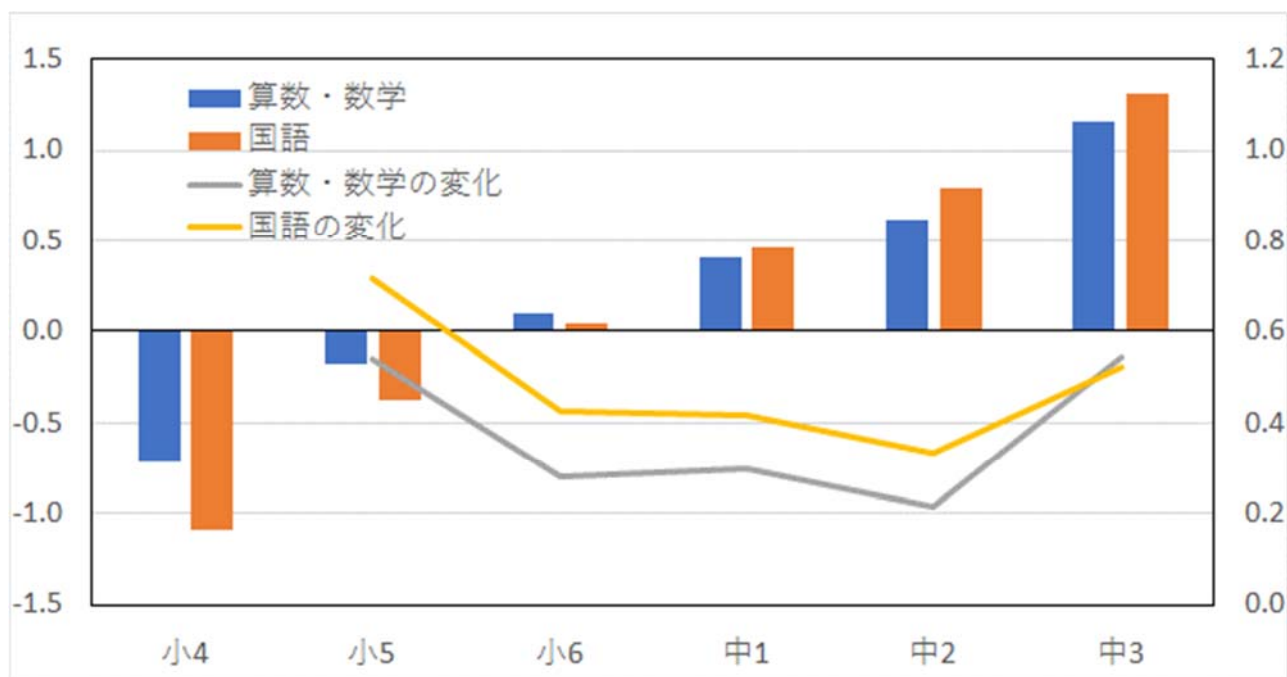
テスト理論に基づく素点方式や偏差値方式などとは異なり、調査時点や学年によらずに生徒の学力を比較することができる。推定された学力は-5.8 から 5.8 の値を取り、生徒の学年が上がるにつれて上昇していく傾向が見て取れる（図 1）。また、この調査の分析は、研究者に対して研究公募を行い、研究のためのデータ利用が認められている（なお、本稿の執筆者は全員この公募研究を受託した研究者である）。自治体が収集したデータを適切に研究者向けに開示し、信頼性の高い研究を行うことで、実際の教育政策や制度設計に役立てるようになることこそ、「エビデンスに基づく教育政策」への第一歩に他ならない。

2-2. 学力の推計値の特徴

まずは、埼玉県学力・学習状況調査の最大の特徴である、IRT に基づく学力の推計値について見ていこう（図 1）。本調査で計測している学力の水準と変化をみると、学年とともに学力の水準は上昇しており、これは科目によらず同じ傾向が見て取れる。一方で、学力の変化をみると、小 5 から中 2 にかけて変化幅が小さくなっていき、中 2 の国語では変化幅がほとんどゼロに近くなっている。一方、中 3 になると、再び変化幅が大きくプラスに転じるが、これは高校受験のために学習量を大幅に増やすためであると考えられる。ただ、こうした学力の変化は、学校の中での学習の質・量のみならず、学習塾での学習時間に大きく影響を受けている（調査票からの推計では、中 2 の学習塾での平均学習時間／週：3.14 時間→中 3 同：4.23 時間／週）。標準偏差も中 2 までは縮小していくが、中 3 になって再び拡

大している。

図 1：各学年の学力の変化



次に、ランダム切片モデルのヌルモデル（説明変数を持たない切片のみのモデル）を用いて、級内相関（Interclass Correlation Coefficient: 以下 ICC）を推定することで、全体の分散に占める学校の分散が占める割合をみる。ICC が高ければ高いほど、学校要因が学力に与える影響が大きいと解釈できるが、国語の学力を被説明変数としたヌルモデルでは、ICC は 3.1～4.1% で、算数・数学でも 3.1～4.3% に過ぎない。このため、個人の学力分散のほとんどが、生徒個人の属性（教育生産関数の X, SES, I などである）によって説明されることになり、日本の学校教育の均質性を窺わせる。この点では、マルチレベルモデリングを用いて、教育生産関数を推定した過去の研究が示す結果とも整合的である（山田, 2009 など）。

表 4：学力の水準（2016 年度）

	学年	N	平均	中央値	標準偏差	ICC
算数・数学	小4	50,227	-0.715	-0.806	1.422	
	小5	48,869	-0.176	-0.326	1.381	0.031
	小6	50,060	0.101	0.010	1.365	0.038
	中1	48,718	0.401	0.281	1.294	0.033
	中2	48,444	0.613	0.599	1.237	0.039
	中3	49,032	1.156	1.069	1.471	0.043
国語	小4	50,225	-1.095	-1.150	1.691	
	小5	48,862	-0.381	-0.492	1.542	0.041
	小6	50,057	0.043	0.014	1.126	0.036
	中1	48,717	0.458	0.595	1.170	0.031
	中2	48,441	0.791	0.809	1.083	0.034
	中3	49,025	1.312	1.252	1.261	0.034
英語	中2	48,445	0.013	-0.007	1.141	0.062
	中3	49,035	0.581	0.445	1.298	0.057

表 5：学力の変化（2015→2016）

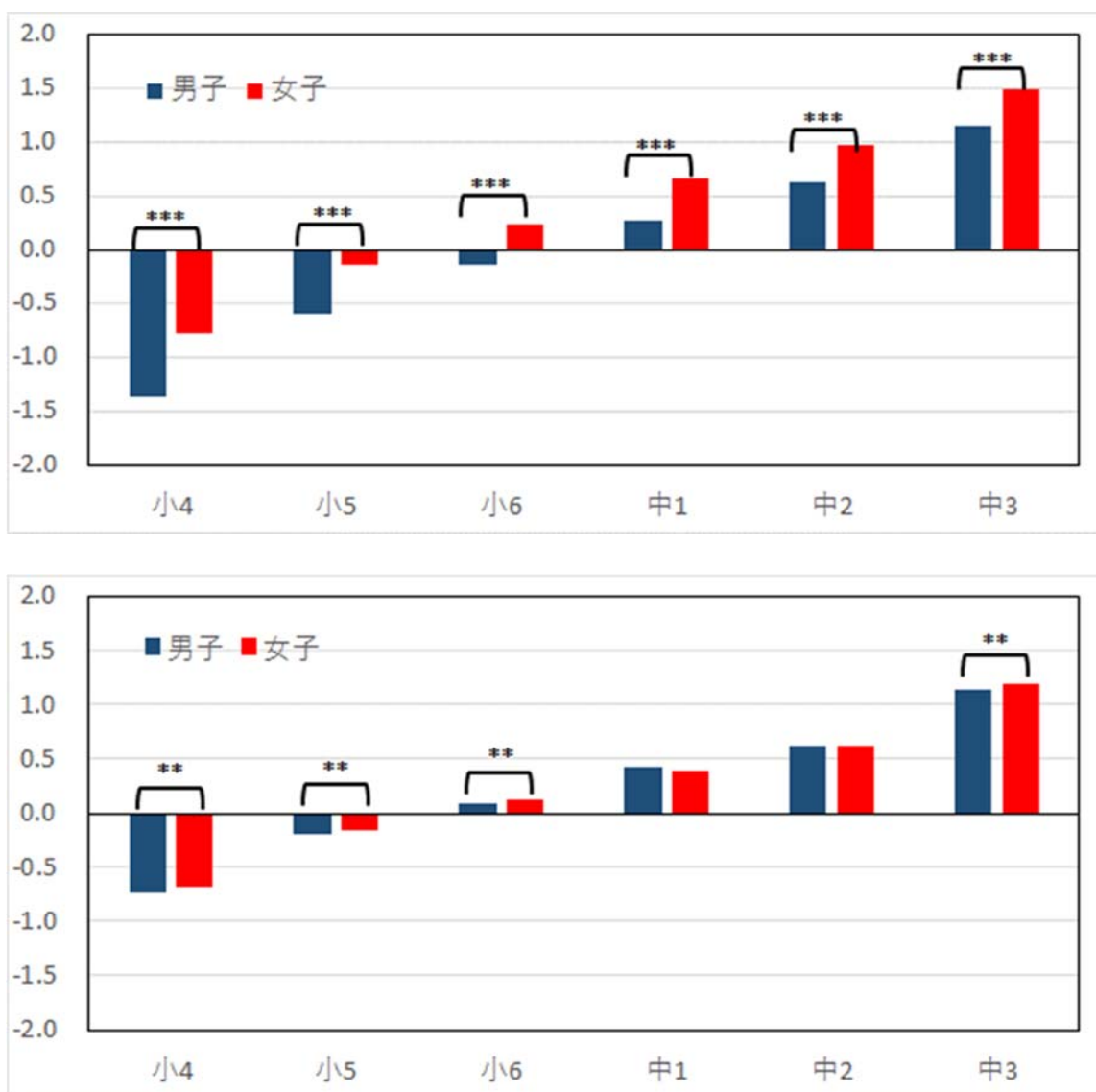
	学年	N	平均	中央値	標準偏差	ICC
算数・数学 (2015→2016)	小5	47,790	0.671	0.614	1.163	0.023
	小6	49,160	0.382	0.329	1.022	0.029
	中1	47,510	0.408	0.379	0.853	0.020
	中2	47,924	0.268	0.265	0.884	0.031
	中3	48,183	0.630	0.591	0.948	0.024
国語 (2015→2016)	小5	47,792	1.128	1.174	1.282	0.032
	小6	49,152	0.401	0.506	1.329	0.024
	中1	47,511	0.382	0.408	1.028	0.015
	中2	47,924	-0.002	0.066	1.087	0.016
	中3	48,174	0.296	0.324	0.960	0.010
英語 (2015→2016)	中3	48,184	0.239	0.373	1.071	0.020

次に、性別による差をみる(図 2)¹⁸。総じて見れば科目によらず、女子の学力が高く、その差は統計的に有意である。PISA や TIMSS、全国学力・学習状況調査のデータを用いた過去の研究では、読解力は女子のほうが高く、理数系科目は男子のほうが高いものの統計的な

¹⁸ ここでは性別として男女とは別に、無回答・無記入の生徒という区別を設け（性別の無回答・無記入は全体の約 0.5%）、グラフでは男子と女子の比較を行っている。

有意差は観察されないことが報告されているが（国立教育政策研究所編, 2008; 妹尾・北條, 2016）、埼玉県のパデータでは、国語、算数・数学ともに女子のほうが高い。数学については中学 1～2 年で男子のほうが高くなっているが、この差は統計的に有意ではなく、中 3 になると再び女子が逆転して、女子のほうが高くなる。

図 2：学力の男女差（上が国語、下が算数・数学、2016 年度）



しかし、文部科学省の「学校基本調査」（2016 年度）によると、埼玉県の 4 年制大学への

進学率は、男子が 55.4%であるのに対して女子が 48.8%にとどまっている上、理数系学部への進学者は男子に比べて極端に少ない。学力的に決して劣るわけではないにもかかわらず、女子が大学進学や理数系学部への進学に消極的なことを考えれば、女子の進学や進路選択には、学力以外の要因が影響している可能性があり、これについては今後さらなる研究が必要とされる。

2－3. 不完備なパネルデータ

このデータを見る上で注意すべきことは、パネルデータが不完備になっていることである。2015 年度調査には参加しているが、2016 年度調査から脱落している生徒は、表 6 で示されたとおり、各学年約 2～3%程度存在している。これは全国学力・学習状況調査における不参加者の割合と同程度であり、病欠や保護者の転勤などに伴って埼玉県外の学校に転校した可能性が考えられる¹⁹。一方、小 6 においてはこの割合は大きく増大し、約 8%にのぼる。表 6 でも明らかなとおり、この 8%の学力は極めて高く、私立中学へ進学した生徒たちである可能性が高い²⁰。他の学年と同程度の生徒の病欠や転校による不参加が存在すると仮定す

¹⁹ 2016 年度「全国学力・学習状況調査」のうち、埼玉県の結果から求められる調査不参加者の割合は本調査と概ね一致する。

(http://www.nier.go.jp/16chousakekkahoukoku/factsheet/16prefecture/11_saitama/index.html)

²⁰ 私立中学へ進学したと見られる生徒のグループはそれ以外の生徒と比較して、IRT の学力推計値が約 1 程度も異なっている。この調査における IRT の学力の推計値の 1 は、古典的テスト理論で推計された偏差値の 10（1 標準偏差）に相当する。

ると、それらを除く約 5～6%が中学進学時に私立中学に進学することで、調査の対象から外れているものと考えられる。

表 6： 2016 年調査不参加者の属性

	学年	2016年の参加	人数	学力平均	学力中央値	標準偏差
算数・数学	小4→小5	あり	47,787	-0.637	-0.613	1.132
		なし	1,046	-1.059	-0.902	1.402
	小5→小6	あり	49,139	-0.137	-0.066	1.094
		なし	965	-0.446	-0.428	1.368
	小6→中1	あり	47,493	-0.053	-0.125	1.055
		なし	3,840	0.926	0.974	1.502
	中1→中2	あり	47,925	0.212	0.136	1.098
		なし	1,480	-0.427	-0.330	1.239
	中2→中3	あり	48,226	0.197	0.153	1.160
		なし	1,258	-0.501	-0.557	1.250
国語	小4→小5	あり	47,793	-0.954	-0.985	1.454
		なし	1,047	-1.387	-1.298	1.811
	小5→小6	あり	49,137	-0.286	-0.298	1.330
		なし	965	-0.640	-0.562	1.495
	小6→中1	あり	47,496	-0.024	-0.039	1.092
		なし	3,839	0.877	0.918	1.512
	中1→中2	あり	47,928	0.350	0.318	1.107
		なし	1,478	-0.161	-0.126	1.231
	中2→中3	あり	48,232	0.436	0.495	0.937
		なし	1,253	0.009	0.131	1.048
英語	中2→中3	あり	48,228	0.038	-0.037	1.067
		なし	1,254	-0.507	-0.629	1.083

2016 年度の調査不参加者が、病欠や転校など予想できない理由によりランダムに欠測している場合は、仮にパネルデータが不完備だった場合も、加重最小二乗法などを利用してサンプルサイズに応じたウエイト付けをすれば問題は生じないが、ランダムに欠測していない場合、サンプルセレクション・バイアスが生じる可能性がある。

データをみると、2016 年度調査に不参加だった生徒の学力は、科目や学年を問わず、それ以外の生徒と比較して統計的に有意に低い傾向があり、ランダムに欠測しているとは言えない状況である。この理由は明確ではないが、病気や不登校による長期欠席や、教員や友人との人間関係がうまく行かずあえて転校したという可能性もあろう。このため、本調査を用いて教育生産関数を推定する際には、脱落サンプル問題について十分な配慮が必要である。

2－4．質問紙調査（非認知能力、学習方略、指導、保護者の社会経済的地位）

埼玉県学力・学習状況調査のもう 1 つの大きな特徴は、豊富な情報を含む質問紙調査である。教育生産関数の主要な投入要素である、生徒本人の属性（ X ）の情報に加え、保護者の社会経済的地位（ SES ）教員の質（ TQ ）、学校資源（ S ）、ピア効果（ P ）などの情報も含まれている。

特に 2016 年度からは、生徒の非認知能力を計測する質問項目が加わっている。近年、経済学では、学力テストや IQ テストで計測できないような非認知能力が教育や労働市場における成果に与える影響が大きいことを明らかにする研究が相次いで発表されている（Heckman & Rubinstein, 2001）。特に教育については、Heckman et al (2006)が、学力テストの結果には非認知能力が影響していることを示しており、特に Heckman et al (2010) は自制心、忍耐力、自己効力感といった非認知能力が、教育や労働市場での成果に与える影響

が大きいことを指摘している。

このため、埼玉県学力・学習状況調査では、質問紙調査の中で、「自制心」(小4・中1)、
「自己効力感」(小5・中2)、「忍耐力」(小6・中3)を()内の学年を対象にして計測し
ている。自制心とは自分の意思で感情や欲望をコントロールすることができること、自己効
力感とは自分自身に対する有能感や信頼感があること、忍耐力とは物事を粘り強く続けてい
くことができる力があること、である。この質問は、一般に心理学で心理現象を測定する方
法として用いられる「心理測定尺度」といわれる手法で、Duckworth, et al (2007)や
Tsukayama, et al (2013)が開発した自制心、忍耐力、自己効力感などの心理測定尺度を、日
本語に翻訳した後、国内の調査を経て、妥当性が検証されているものである。実際に調査で
用いられた質問項目が表 7 のようになっている。

こうした心理測定尺度は、心理現象の厚生概念が反映されるように複数の質問項目からな
っているのが一般的である。例えば、自制心であれば、「授業で必要なものをわすれた」「ほ
かの子たちが話をしている時に、その子たちのじゃまをした」などの 8 問について、5 件法
(1: まったくあてはまらない, 2: あまりあてはまらない, 3: すこしあてはまる, 4: よくあ
てはまる, 5: とてもよくあてはまる) で回答してもらい、この 8 問の回答を足し上げると
自制心を数字で計測することができるというわけである²¹。

²¹ ただし、質問票の中には、「逆転項目」が含まれていることがある。例えば、「勤勉性」の質問項目に

表 7：非認知能力の質問項目

自制心

授業で必要なものをわすれた
 ほかの子たちが話をしているときに、その子たちのじゃまをした
 何からんぼうなことを言った
 机・ロッカー・部屋が散らかっていたので、必要なものを見つけることができなかった
 家や学校でカッとなってキレた
 先生が、自分にたいして言っていたことを思い出すことができなかった
 きちんと話を聞かないといけないときにぼんやりしていた
 イライラしているときに、先生や親に口答えをした

自己効力感

授業ではよい評価をもらえるだろうと信じている
 教科書の中でいちばんむずかしい問題も理解できると思う
 授業で教えてもらった基本的なことは理解できたと思う
 先生が出したいちばんむずかしい問題も理解できると思う
 学校の宿題や試験でよい成績をとることができると思う
 学校でよい成績をとることができるだろうと思う
 授業で教えてもらったことは使いこなせると思う
 自分はこの授業でよくやっているほうだと思う

勤勉性

うっかりまちがえたりミスしたりしないように、やるべきことをやります
 ものごとは楽しみながらがんばってやります
 自分がやるべきことにはきちんとかかわります
 授業中は自分がやっていることに集中します
 宿題が終わったとき、ちゃんとできたかどうか何度も確認をします
 ルールや順番は守ります
 だれかと約束をしたら、それを守ります
 じぶんの部屋やつくえのまわりはちらかっています
 何かをはじめたら、ぜったい終わらせなければいけません
 学校で使うものはきちんと整理しておくほうです
 宿題を終わらせてから、遊びます
 気がちってしまうことはあまりありません
 やらないといけないことはきちんとやります

は、「じぶんの部屋やつくえのまわりはちらかっています」という質問があり、他の質問が「当てはまる」場合に勤勉性が高くなる質問なのに対して、この質問だけは測定に向きが逆になっているため、これは逆転項目である。逆転項目の質問に対する点数は反転させて、足し上げる必要がある。

表 8：非認知能力（2016 年）

		N	平均	中央値	標準偏差	ICC
自制心	小4	50,184	31.665	33	6.194	
自制心	中1	48,709	31.138	32	5.966	0.035
自己効力感	小5	48,865	28.245	29	6.295	0.028
自己効力感	中2	48,444	24.150	24	6.547	0.016
勤勉性	小6	50,011	48.657	50	8.521	0.026
勤勉性	中3	49,046	46.532	47	8.036	0.019

表 8 をみると、非認知能力を被説明変数としたヌルモデルでは、ICC は 1.6～3.5%で、個人の非認知能力の分散も、学校ではなく個人の要因によって説明される部分が多い。

小学生よりも中学生において平均値が低くなっているものの、同じ個人の変化をみているのではないため、非認知能力が低下しているとまでは言えず、今後どのように変化していくかを観察していく必要がある。

また、これ以外に学習の効果を高めることをめざして生徒が意図的に行う活動である「学習方略」についても全学年で計測している。生徒の学習方略は、教員からの指導を反映していると考えられており、教員の質（**TQ**）の代理変数として用いることができる。この学習方略は、細かくみると柔軟的方略（学習の進め方を自分の状態に合わせて柔軟に変更していく方略）、プランニング方略（計画的に学習に取り組もうとする方略）、作業方略（ノートに書いたり、声に出したりといった、「作業」を中心として学習を進める方略）、人的リソース方略（友人を利用して学習を進める方略）、認知的方略（理解や精緻化、集中力と言った認知的な働きを重視して学習を進める方略）、努力調整方略（「苦手」などの感情をコントロール

して学習への動機を高める方略) の6つのカテゴリにわかれている²²。実際に調査で用いられた質問項目が表9のようになっている。

表9：学習方略の質問項目

柔軟的方略	勉強のやり方が、自分にあっているかどうかを考えながら勉強する
	勉強でわからないところがあったら、勉強のやり方をいろいろ変えてみる
	勉強しているときに、やった内容をおぼえているかどうかをたしかめる
	勉強する前に、これから何を勉強しなければならないかについて考える
プランニング方略	勉強するときは、さいしょに計画をたててからはじめる
	勉強をしているときに、やっていることが正しくできているかどうかをたしかめる
	勉強するときは、自分できめた計画にそってこなす
作業方略	勉強しているとき、たまに止まって、一度やったところを見なおす
	勉強するときは、参考書や事典などがすぐ使えるように準備しておく
	勉強する前に、勉強に必要な本などを用意してから勉強するようにしている
	勉強していて大切だと思ったところは、言われなくてもノートにまとめる
人的リソース方略	勉強で大切なところは、くり返して書いたりしておぼえる
	勉強でわからないところがあったら、友達にその答えをきく
	勉強でわからないところがあったら、友達に勉強のやり方をきく
	勉強のできる友達と、同じやり方で勉強する
認知的方略	勉強するときは、最後に友達と答えあわせをするようにする
	勉強するときは、内容を頭に思い浮かべながら考える
	勉強をするときは、内容を自分の知っている言葉で理解するようにする
	勉強していてわからないところがあったら、先生にきく
努力調整方略	新しいことを勉強するとき、今までに勉強したことと関係があるかどうかを考えながら勉強する
	学校の勉強をしているとき、とてもめんどくさいと思うことがよくあるので、やろうとしていたことを終える前にやめてしまう
	いまやっていることが気に入らなかったとして、学校の勉強でよい成績をとるためにいっしょうけんめいがんばる
	授業の内容がむずかしいときは、やらずにあきらめるか簡単なところだけ勉強する
	問題が退屈でつまらないときでも、それが終わるまでなんとかやりつづけられるように努力する

²² Pintrich & de Groot(1990)を参考にした。

表 10：学習方略（2016 年）

		N	平均	中央値	標準偏差	ICC
全体	小4	48,856	73.515	75	14.885	
	小5	47,982	73.526	75	15.171	0.042
	小6	49,506	73.904	75	15.392	0.045
	中1	48,147	75.317	77	14.418	0.059
	中2	47,960	70.291	72	15.719	0.030
	中3	48,641	68.878	70	15.522	0.023
努力調整方略	小4	48,846	16.319	17	3.078	
	小5	47,980	16.184	17	3.040	0.018
	小6	49,499	15.941	16	3.056	0.025
	中1	48,146	15.985	16	2.927	0.029
	中2	47,958	14.588	15	3.364	0.017
	中3	48,641	14.224	14	3.395	0.013
作業方略	小4	49,430	13.701	14	3.742	
	小5	48,398	13.669	14	3.725	0.036
	小6	49,881	13.882	14	3.713	0.035
	中1	48,638	14.492	15	3.530	0.044
	中2	48,432	14.081	14	3.614	0.024
	中3	49,105	13.949	14	3.585	0.020
認知的方略	小4	49,523	15.593	16	3.105	
	小5	48,441	15.541	16	3.079	0.030
	小6	49,912	15.499	16	3.060	0.031
	中1	48,647	15.461	16	2.849	0.034
	中2	48,435	14.407	15	3.195	0.021
	中3	49,104	14.329	15	3.202	0.018
プランニング方略	小4	49,307	14.265	15	3.584	
	小5	48,360	14.307	15	3.575	0.029
	小6	49,864	14.315	15	3.550	0.033
	中1	48,631	14.599	15	3.325	0.039
	中2	48,425	13.718	14	3.495	0.020
	中3	49,102	13.379	14	3.462	0.016
人的リソース方略	小4	49,290	11.627	12	3.679	
	小5	48,352	11.873	12	3.583	0.036
	小6	49,867	12.160	12	3.544	0.035
	中1	48,626	12.444	13	3.429	0.034
	中2	48,430	11.874	12	3.624	0.022
	中3	49,101	11.518	12	3.631	0.018
柔軟的方略	小4	49,421	14.176	14	3.563	
	小5	48,401	14.038	14	3.546	0.035
	小6	49,882	14.129	14	3.528	0.036
	中1	48,636	14.330	15	3.307	0.050
	中2	48,432	13.619	14	3.491	0.022
	中3	49,100	13.461	14	3.447	0.016

2-5. 指導

また、埼玉県学力・学習状況調査では、学習方略以外に、教員の指導方法について子どもの受けとめを尋ねた質問項目がある。特に、各科目で「授業で課題を解決するときにみんなでいろいろな考えを発表することはありませんか？」などのような質問を行うことで、間接的に学級内で教員がどのような指導を行ったかを推し量ることができる。これも教員の質（*TQ*）の代理変数である。特に、本調査では、「アクティブラーニング」（＝主体的かつグループディスカッションやディベートなどを取り入れた参加型の学び）と呼ばれる指導が行われているかどうかを、生徒側の経験から汲み取ろうとしている²³。

アクティブラーニングは、文部科学省が公表した学習指導要領改定案の中でも「主体的・対話的で深い学び」と呼ばれ、中央教育審議会の答申においてもキーワードの1つとして掲げられている。近年の経済学の研究の中には、教員の学歴や経験年数、免許の有無などのような「履歴書に掲載されるような」属性は教員の質を適切にあらわしておらず、むしろ教員の教室での実践の重要性を強調する研究もある（e.g., Aslam & Kingdon, 2011）。しかし、学

²³ ただし、この質問項目の中には主体的な学びの一環として「ドリル学習」なども含まれていることから、「アクティブラーニング」というよりは、「授業充実度」あるいは「多様な学び」と定義するほうが妥当かもしれない。また、繰り返しになるが、この質問項目は、生徒がどのような指導を受けていたかを尋ねたものであり、実際に教員がどのような指導をしたかをあらわしていない。このため、従来型の授業と比べてアクティブラーニングのほうが効果があったと結論付けられるものではないことに注意が必要である。

力を上げるという目的に対して、アクティブラーニングのほうが、いわゆる「知識詰め込み型」として評判の悪い従来の指導法よりも優れているというエビデンスは多くない。例えば、Bietenbeck (2014)は、TIMSS のデータを用いて、従来型の指導法は、定型化された問題を解く能力を高めるものの、論理的思考力を高めるわけではないことを発見し、一方、アクティブラーニングはこれと全く逆になることを発見している。つまり、学力テストが計測している教育成果が、論理的思考力の獲得量を計測するものになっていなければ、アクティブラーニングは学力テストに負に影響しているという結論になると述べている。ここではまず、表 11 のような質問項目を足し上げることで、各科目ごとのアクティブラーニング指導の指標を作成し、次章で学力や非認知能力、学習方略との関係を見ていくことにする。

表 11：指導に関する質問項目

教師との関係性	学級での生活は楽しかったですか
	学級は落ち着いて学習する様子でしたか
	学校での生活には満足していましたか（運動会や遠足などの学校行事も入ります）
	学校の先生たちは自分の良い所をほめてくれましたか
	学校の先生たちは心配事の相談にのってくれましたか
アクティブラーニング(国語)	友達の考えを聞いて、文章の内容や表現の仕方がよくわかったこと
	自分の考えを理由をつけて発表したり、書いたりできたこと
	ノートやワークシート、プリントに書いた授業のまとめを先生に見てもらうこと
	ドリルなどをする
	グループで活動するときに、一人の考えだけでなくみんなで考えを出し合って課題を解決すること
アクティブラーニング(算数)	授業で課題を解決するときに、みんなで色々な考えを発表すること
	授業のはじめに、先生から、どうしたら課題を解決できるか考えるように言われること
	授業のはじめには気が付かなかった疑問が、授業の終わりに、頭に浮かんできたこと
	課題を解決するときに、それまでに習ったことを思い出して解決出来たこと
	自分の考えを理由をつけて発表したり、書いたりできたこと
アクティブラーニング(英語)	ノートやワークシート、プリントに書いた授業のまとめを先生に見てもらうこと
	ドリルなどをする
	グループで活動するときに、一人の考えだけでなくみんなで考えを出し合って課題を解決すること
	授業で課題を解決するときに、みんなで色々な考えを発表すること
	授業のはじめに、先生から、どうしたら課題を解決できるか考えるように言われること
アクティブラーニング(英語)	授業のはじめには気が付かなかった疑問が、授業の終わりに、頭に浮かんできたこと
	授業で、友達と英語を使って活動することで、新しい英語の表現を使えるようになりましたか
アクティブラーニング(英語)	授業で、自分や友達の考えや気持ちなどについて、英語で聞く、話す、読む、書くなどの活動を行っていましたか
	授業で、英語を使って活動することで、自分を英語を使ってみたいと思うようになりましたか

表 12：指導（2016 年度）

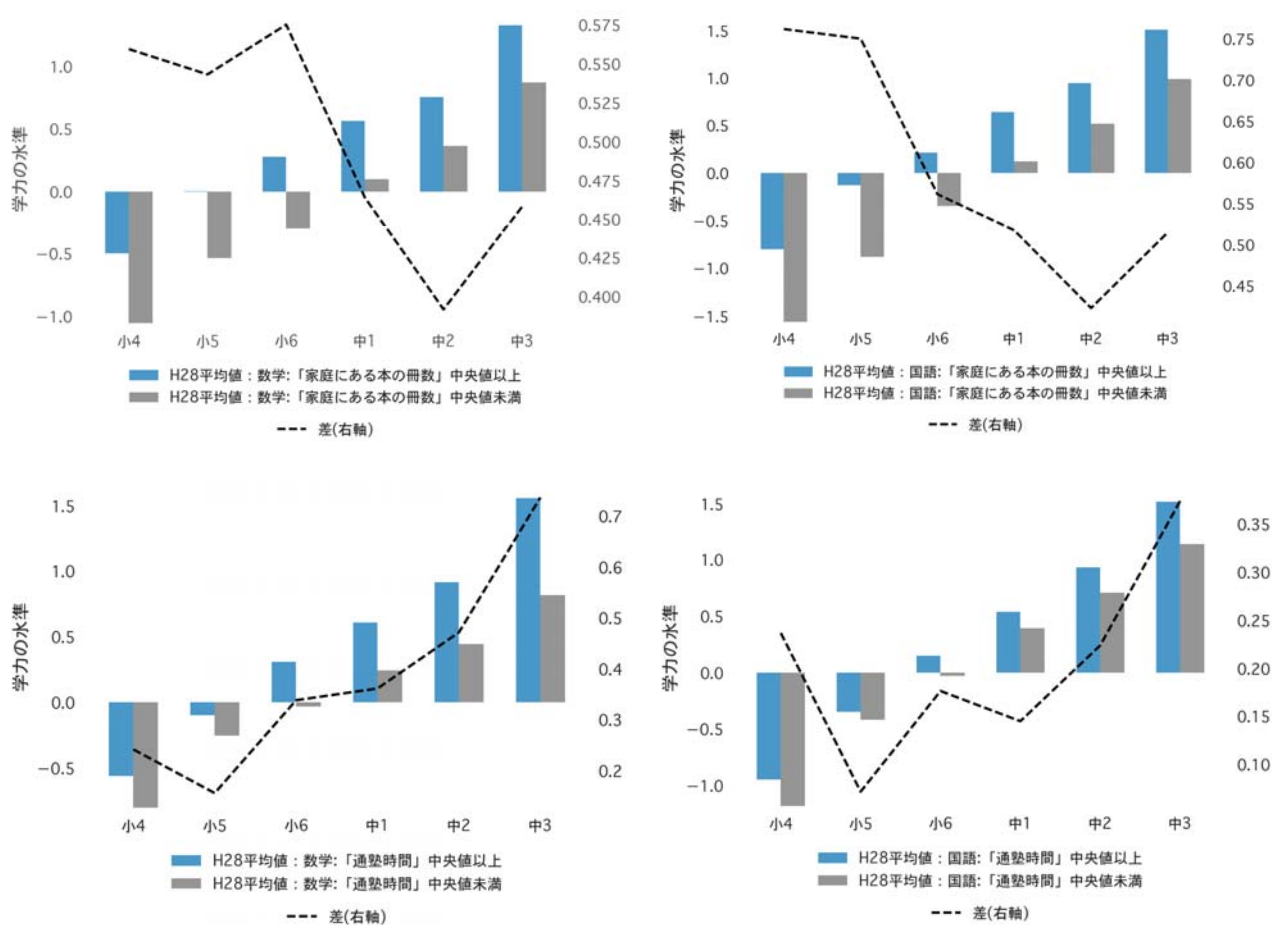
		N	平均	中央値	標準偏差	ICC
アクティブラーニング	小4	50,054	25.823	26	4.385	
(国語)	小6	49,993	26.455	27	3.909	0.042
	中2	48,326	25.042	25	4.147	0.063
アクティブラーニング	小5	48,754	26.401	27	3.981	0.039
(算数・数学)	中1	48,680	26.680	27	3.586	0.054
	中3	48,899	23.740	24	4.521	0.070
アクティブラーニング	中2	48,290	9.782	10	2.215	0.052
(英語)	中3	48,845	9.376	10	2.324	0.059
教師との関係	小4	50,228	22.150	23	2.689	
	小5	48,873	22.111	23	2.774	0.062
	小6	50,061	21.840	22	2.861	0.077
	中1	48,720	22.106	23	2.713	0.084
	中2	48,450	21.320	22	2.875	0.050
	中3	49,078	21.226	22	2.896	0.040

2－6．保護者の社会経済的地位

埼玉県学力・学習状況調査では、保護者の社会経済的地位（*SES*）の代理変数として用いることができるものとして、家庭の文化資本をあらわす家にある蔵書数（「家には、自分や家の人が読む本がどれくらいありますか」を、1:ほとんどない（0 から 10 冊）、2: 本棚 1 列分（11～25 冊）、3: 本棚 1 つ分（26～100 冊）、4: 本棚 2 つ分（101～200 冊）、5: 本棚 3 つ分（201～300 冊）の中から選択）や通塾（「学習塾（家庭教師に教わっている場合もはいります）で 1 週間のうち、どのくらいの時間、勉強しますか」を 1: 通っていない、2: 12 時間以上、3: 10 時間以上 12 時間より少ない、4: 8 時間以上 10 時間より少ない、5: 6 時間以上 8 時間より少ない、6: 4 時間以上 6 時間より少ない、7: 2 時間以上 4 時間より少ない、8: 2 時間より少ない、の中から選択）などが考えられる。

この 2 つの変数の中央値を取り、中央値よりも高いグループを社会経済的地位が高いグループ、中央値よりも低いグループを社会経済的地位が低いグループと定義し、2 つのグループの間で学力を比較してみると、図 3 から明らかなように、2 つのグループには明確な学力格差があり、この差は統計的に有意である。

図 3：保護者の社会経済的地位による学力格差

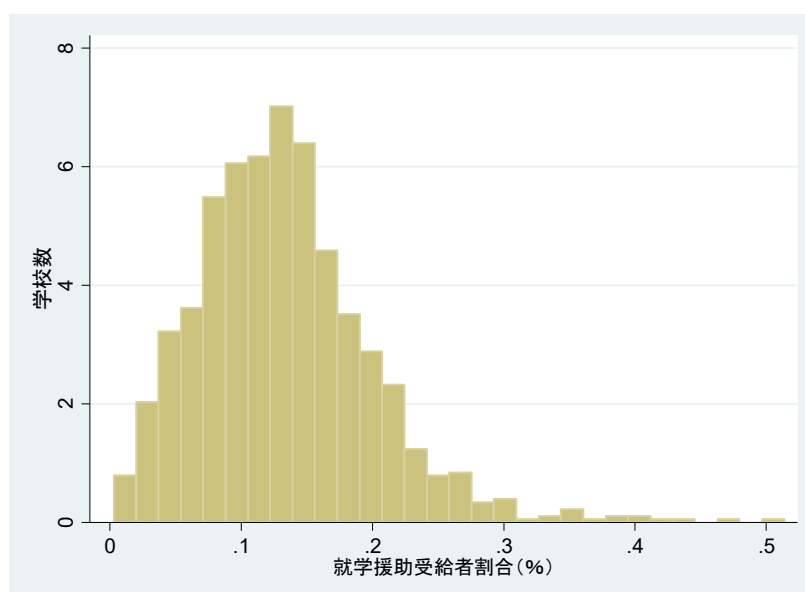


驚くべきことに、保護者の社会階層による格差は、既に小学校 4 年生の時点では顕在化しており、親の社会経済的地位による分断がかなり早い学齢から始まっていることがわかる。この点は、日本のデータを用いた先行研究とも整合的である (Matsuoka et al, 2015 など)。

特に通塾時間については、学年が上がるほどに格差が拡大していく傾向があり、学校外教育へのアクセスが格差を拡大している可能性が示唆される。この意味においては、「貧困の世代間連鎖」の問題を学校の中だけで解決しようとするのは極めて難しく、小学校低学年かあるいはその以前から始まっていると考えられる保護者の社会経済的地位による学力格差を解消していくには、就学前教育の充実や放課後の学校外教育へのアクセスなどの充実も検討していく必要がある。

また、学校質問紙から、就学援助受給者比率をみると(図 4)、埼玉県全体では約 13.2%と全国平均よりは低くなっているものの(平成 25 年度、15.4%)、受給者比率が最大の学校で 51.4%、最小の学校で 0.3%とかなり学校によってばらつきがあり、学校によって抱えている「子どもの貧困」の問題の深刻さには差があることがわかる。

図 4：学校別の就学援助率の分布



この章の最後に、質問紙の改善、拡張について提言を行いたい。埼玉県学力・学習状況調査の質問紙は質問項目が充実しており、計測が困難な非認知能力や学習方略についての質問項目は学術的に見て妥当性の高いものを厳選しているなど、様々な工夫が凝らされ、教育政策の評価や学校における指導・資源配分を明らかにする分析を行うことを念頭に置いた設計になっている。

しかし、この調査を用いて行う分析が、教育政策の評価や学校における指導・資源配分の改善を一義的なものとするのであれば、教員の質（**TQ**）に関する変数は、計測誤差が発生しないよう、教員が直接、指導や資源配分に対して答えたものであるほうが望ましい。また保護者の社会経済的地位（**SES**）についても、保護者が直接答えてたものであるほうが望ましい。文化的、あるいは経済的な資源にとどまらず、親子の関係や生活や学習へのかかわりなどについての情報は、貧困の世代間連鎖を断ち切るための研究蓄積には必須の情報である。

今後、抽出調査にしつつ費用を抑えるなどの方法も視野にいれるなどして、教員や保護者向けの調査が拡充されることが望まれる。

3. 埼玉県学力・学習状況調査を用いた分析

3-1. 教育生産関数の推定

ここでは、埼玉県学力・学習状況調査のデータを用いて、教育生産関数の推定を行う。その際には、2 時点間のパネルデータであることを利用して、2015 年の生徒個人の学力の水準をコントロールした付加価値モデルとして推定する。推計結果は表 13 で示されたとおりである。非認知能力やアクティブラーニングの実施の有無など、収集された情報が学年によって異なっているため、5 つのモデルを、国語と算数・数学のそれぞれで推定した。

結果を見てみると、前年の学力からの影響を推定してみると、国語では、前年の学力の変動が今年度の学力の変動の約 40%を説明しているのに対して、数学（算数）は約 60%を説明している。これは、特に国語に比べると数学・算数は、前年にわからないところがあってもつまづいてしまうと翌年にも大きな影響を残してしまうことを意味しており、数学・算数は特につまづきの解消や復習を行うことが重要であることがわかる。

また、先行研究が指摘するとおり、非認知能力（自制心、自己効力感、勤勉性）と学力との間には正の相関がある。このことは、非認知能力の獲得が認知能力を向上させる可能性を示唆しており、海外での研究成果（Cunha, et al 2010）とも一致する。これ以外にも、アクティブラーニングを実施しているクラスの生徒は科目によらず、学力が高い傾向があること

も示されている。²⁴一方、学習方略については、総じて、算数・数学では正の相関がみられるものの、逆に国語では負になっている。この理由としては、後述するとおり、学習方略を構成する要因のうち、学力に正の影響を持つものと、負の影響を持つものがあり、科目によってばらつきがあるためと考えられる。また、親の社会経済的地位をあらわす変数は、概ね予想したとおりの符号で、一貫して統計的に有意になっている。一方、教師生徒比率は、モデルや科目によって正になっていたり、負になっていたりして一貫しない。就学援助受給比率は、概ね負で統計的に有意になっている。

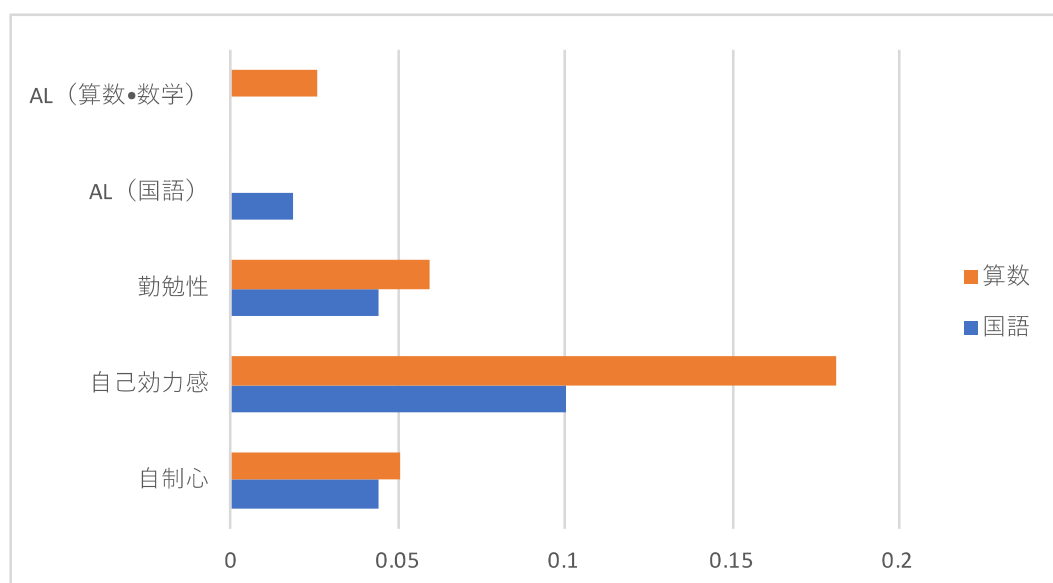
今度はこれについて、非認知能力と指導、学習方略の影響の大きさを比較するために、図 5：標準化係数で標準化係数を見てみると、非認知能力は自己効力感の影響が大きく、指導としては算数へのアクティブラーニングの指導の影響が大きいことがわかる。

²⁴ ここでの分析では表 11「指導に関する質問項目」に記載されている「アクティブラーニング」の中でも、「グループで活動するときに、一人の考えだけでなくみんなで考えを出し合って課題を解決すること」や「授業で課題を解決するときに、みんなで色々な考えを発表すること」などアクティブラーニングと解釈できる項目のみ利用した。

表 13：回帰分析

	国語					算数				
	Model1	Model2	Model3	Model4	Model5	Model1	Model2	Model3	Model4	Model5
2015年度の学力(国語)	0.5352*** (0.0027)	0.4450*** (0.0025)	0.5514*** (0.0027)	0.3924*** (0.0022)	0.6472*** (0.0032)					
2015年度の学力(算数)						0.8735*** (0.0039)	0.6643*** (0.0036)	0.5923*** (0.0041)	0.7242*** (0.0039)	0.8310*** (0.0036)
非認知能力										
自制心	0.0086*** (0.0007)					0.0109*** (0.0007)				
自己効力感		0.0257*** (0.0006)	0.0191*** (0.0008)				0.0406*** (0.0006)	0.0312*** (0.0009)		
勤勉性				0.0072*** (0.0005)	0.0070*** (0.0005)				0.0107*** (0.0006)	0.0111*** (0.0006)
周囲の学力										
2015年度の周囲の学力(国語)	-0.1198*** (0.0205)	-0.0634*** (0.0180)	-0.1695*** (0.0199)	-0.1120*** (0.0169)	-0.2240*** (0.0255)					
2015年度の周囲の学力(算数)						-0.1467*** (0.0299)	-0.1225*** (0.0253)	-0.1385*** (0.0313)	-0.1892*** (0.0323)	-0.2574*** (0.0302)
指導や教員の質										
アクティブラーニングの実施(国語)		0.0197* (0.0115)		0.0443*** (0.0127)			0.0064 (0.0123)		0.0815*** (0.0155)	
アクティブラーニングの実施(算数)	0.0340*** (0.0130)		0.0851*** (0.0173)		0.0122 (0.0108)	0.0465*** (0.0137)		0.0706*** (0.0186)		0.0244** (0.0115)
学習方略	0.0012 (0.0018)	-0.0033** (0.0015)	-0.0059*** (0.0022)	-0.0043*** (0.0017)	-0.0013 (0.0017)	0.0017 (0.0019)	-0.0018 (0.0016)	-0.0060*** (0.0023)	-0.0012 (0.0021)	-0.0009 (0.0018)
家庭の社会経済的地位										
通塾時間（週あたり）	0.0039* (0.0020)	0.0061*** (0.0018)	-0.0022 (0.0023)	0.0291*** (0.0018)	0.0292*** (0.0019)	0.0504*** (0.0021)	0.0461*** (0.0020)	0.0200*** (0.0025)	0.0501*** (0.0022)	0.0719*** (0.0020)
家庭にある本の冊数	0.0610*** (0.0033)	0.0345*** (0.0029)	0.0766*** (0.0044)	0.0655*** (0.0034)	0.0605*** (0.0033)	0.0403*** (0.0034)	0.0222*** (0.0031)	0.0571*** (0.0046)	0.0596*** (0.0041)	0.0309*** (0.0035)
学校の資源										
習熟度別授業の実施（国語）		0.1206*** (0.0367)	0.2361*** (0.0742)	0.008 (0.0631)	0.1393*** (0.0448)					
習熟度別授業の実施（算数）							0.0145 (0.0284)	0.2056** (0.0912)	0.1751** (0.0758)	-0.3342*** (0.0362)
就学援助率	-0.7201*** (0.2019)	-0.7810*** (0.1423)	-0.7775*** (0.1755)	-0.1776 (0.1367)	-0.9031*** (0.2140)	-0.3876* (0.2121)	-0.1801 (0.2470)	-0.2258*** (0.0808)	0.2654 (0.2505)	-3.1209*** (0.3230)
家庭にある本の冊数	-1.5271*** (0.5184)	0.05 (0.0806)	-1.6464*** (0.4733)	0.017 (0.3682)	0.2674*** (0.1034)	0.0653 (0.5446)	-0.0244 (0.1137)	-0.7825*** (0.2515)	1.2395** (0.5601)	1.0401*** (0.1438)
性別	○	○	○	○	○	○	○	○	○	○
生まれ月	○	○	○	○	○	○	○	○	○	○
学年	○	○	○	○	○	○	○	○	○	○
学校固定効果	○	○	○	○	○	○	○	○	○	○
サンプルサイズ	45341	45792	45212	46662	45757	45339	45824	45210	46669	45956
決定係数	0.54	0.54	0.6	0.52	0.54	0.58	0.6	0.43	0.51	0.62

図 5：標準化係数



3－2．学習方略と学力の関係

表 13 の分析では、学習方略の係数の符号が安定せず、科目によってもばらつきが見られた。そこで、学習方略を構成する要因のうち、特に学力と相関している要因を特定するために、表 14 のような下記分析を実施した。これは表 13：回帰分析 と同様の回帰分析の中で、学習方略だけを下記のように 6 つに分解したものである。なお、表の簡略化のために学習方略に係る部分のみ表示している。

表 14：学習方略と学力の関係

	H28平均値：数学	H28平均値：国語
柔軟的方略	-0.0001 (0.001)	-0.005*** (0.001)
プランニング方略	0.002*** (0.001)	0.001* (0.001)
作業方略	-0.002** (0.001)	-0.003*** (0.001)
人的リソース方略	-0.026*** (0.001)	-0.025*** (0.001)
認知的方略	0.023*** (0.001)	0.023*** (0.001)
努力調整方略	0.049*** (0.001)	0.036*** (0.001)
Observations	235765	235724

表 14 を見ると、学習方略の中には、学力と正の相関関係にある方略と、逆に負の関係がある方略がある。プランニング方略、認知的方略、努力調整方略は科目によらず、正に統計的に有意な関係がある。しかし、作業方略や人的リソース方略は負に統計的に有意であり、特に国語に関しては柔軟的方略も負に統計的に有意である。この意味では、生徒に学力と正の相関があるプランニング方略、認知的方略、努力調整方略などを身に付けさせることが重要である可能性がある。つまり、生徒に計画的に学習に取り組ませること、理解や精緻化を図ること、「苦手」などの感情をコントロールさせることが重要だということである。

3-3. アクティブラーニングによる格差縮小の可能性

アクティブラーニングと学力向上の関連について、先行研究（前馬, 2016; 須藤, 2013; など）を参考に、学習機会の観点で追加分析を行った。表 15 にあるように、数学については、アクティブラーニングと非通塾者の交互作用が有意になっている。これは、通塾者よりも非通塾者のほうがアクティブラーニングによる便益をより大きく受けていると解釈できる。公立学校内の授業にだけ学力向上機会を依拠する非通塾層にとって特に便益があるということとは、数学については、アクティブラーニングによって学校外学習機会格差に基づく学力格差を縮小する可能性を示唆している。

ただ、アクティブラーニングの数学と国語それぞれに対する主効果、それに数学について

の非通塾者に対するアクティブラーニングのより大きな便益は、他の指導手法よりもアクティブラーニングのほうが強い効果を持つ、という意味ではない。アクティブラーニングの計測そのものの難しさに加えて生徒回答による指標化であることから、結果に留意する必要がある。

一方、IRT に基づく前年度学力やその他の多くの変数を統制した上でも、アクティブラーニングと解釈できる授業を前年度に受けた生徒の数学と国語の学力は向上し、数学については非通塾層の学力がより上がったことは、公教育が（指導の方法によっては）学校外学習機会格差による学力格差を是正する機能を持つことを示していることから、注目に値する結果といえる。今後もさらなるデータ蓄積によって、学習機会格差に対して、学校に何ができるのか明らかにしていく必要がある。

表 15：アクティブラーニングと非通塾の交互作用

	H28平均値:数学	H28平均値:国語
アクティブラーニング	0.0235*** (0.0071)	0.0362*** (0.0087)
非通塾×アクティブラーニング	0.0683*** (0.0065)	0.0005 (0.0089)
コントロール		
前年度学力	○	○
通塾の有無	○	○
本の冊数	○	○
性別	○	○
生まれ月	○	○
学年	○	○
前年度学力クラス平均	○	○
通塾前年度クラス平均	○	○
本の冊数前年度クラス平均	○	○
就学援助率	○	○
生徒教師比率	○	○
学校固定効果	○	○
サンプルサイズ	136816	92799
決定係数	0.59	0.56

3-4. アクティブラーニングが非認知能力に与える影響

非認知能力が学力と正の相関関係があることは既を示したが、次に、非認知能力を被説明変数として非認知能力の決定要因を明らかにすることを試みる。Cunha et al (2010)は、非認知能力の獲得量は年齢に依存せず、非認知能力は認知能力を向上させるが、その逆は観察されないことを明らかにしている。

表 16 で結果をみると、科目によらず、アクティブラーニングの実施は、全ての非認

知能力と正の相関があることがわかる。また、この回帰分析では、「教員との関係」をあらわす変数を作成し、それを加えた。変数の作成については表 11 を参照してほしいが、主に「教員が自分の良いところを褒めてくれたか」「心配ごとの相談に乗ってくれたか」などの質問項目からなっている。この教員との関係も、全ての非認知能力と正の相関があることがわかる。保護者の社会経済的地位もまた、非認知能力に影響を与えていることがわかる一方で、学校資源は非認知能力には影響を与えない。

表 16 : 非認知能力と指導

	自制心		自己効力感			勤勉性		
	Model 1	Model 2	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
アクティブラーニングの実施(国語)			0.2352*** (0.0072)			0.3295*** (0.0100)		
アクティブラーニングの実施(算数・数学)	0.2452*** (0.0085)			0.4197*** (0.0072)			0.1418*** (0.0081)	
教員との関係		0.4467*** (0.0098)			0.2608*** (0.0067)			0.3182*** (0.0083)
学習方略	0.1555*** (0.0021)	0.1617*** (0.0019)	0.2021*** (0.0019)	0.1821*** (0.0019)	0.2217*** (0.0012)	0.2693*** (0.0025)	0.2725*** (0.0023)	0.2843*** (0.0016)
塾に通っている時間	0.0336** (0.0133)	0.0462*** (0.0131)	0.3357*** (0.0128)	0.2336*** (0.0112)	0.2893*** (0.0086)	0.0260* (0.0155)	-0.0601*** (0.0142)	-0.0093 (0.0105)
家庭にある本の冊数	-0.0676*** (0.0208)	0.0153 (0.0205)	0.3102*** (0.0204)	0.3109*** (0.0206)	0.4108*** (0.0147)	0.1484*** (0.0282)	0.0417* (0.0248)	0.1774*** (0.0186)
教師生徒比率	-0.9801 (3.3487)	-2.4799 (3.3036)	-0.2774 (0.9542)	-1.1335 (3.1865)	0.456 (3.1181)	-2.5342 (4.5180)	2.0179 (1.2775)	-2.133 (4.0781)
就学援助率	-0.5345 (1.3015)	-1.1989 (1.2839)	0.9545 (2.1081)	0.5385 (1.2204)	1.082 (1.0781)	-0.4391 (1.7602)	2.5725 (2.9093)	-0.7592 (1.4332)
性別	○	○	○	○	○	○	○	○
学年	○	○	○	○	○	○	○	○
生まれ月	○	○	○	○	○	○	○	○
学校固定効果	○	○	○	○	○	○	○	○
サンプルサイズ	44,821	44,821	45,379	44,454	89,837	46,336	45,554	91,896
決定係数	0.28	0.30	0.38	0.44	0.45	0.40	0.37	0.40

4. 埼玉県戸田市で実施した教員調査の分析

4-1. 教員調査の概要

上記述べたとおり、埼玉県学力・学習状況調査では、生徒、学校、市町村を対象とした質問紙調査を実施しているが、教員と保護者を対象とした調査を実施していない。このため、本調査チームは、教員の指導と学力の関係について追加的な分析を行う目的で、2016 年度に埼玉県戸田市の教員を対象とした調査を実施した。教員の非認知能力、学校の運営体制、教員の協同体制、家庭学習への手立て、アクティブラーニングなどの指導法の選択、保護者との連携など多岐にわたる項目について質問紙による調査である。

図 6 と図 7 で学校の運営体制（e.g, 学校として学力向上に向けた課題や目標が提示されたか、などの 6 項目に 4 件法で回答）や教員の協働体制（e.g., 教職員同士がお互いに助け合う協力的な雰囲気があったか、などの 6 項目に 4 件法で回答）といった組織のあり方やマネジメントへの評価を見てみると、概ね小学校のほうが肯定的な評価をしているように見受けられる。また、家庭学習については、小学校教員は数名を除いて全員の教員が計画的に宿題を与え、家庭学習を促したと答えているのに対して、中学校ではばらつきがある状況となっている。具体的な指導法について、アクティブラーニングの実施は（e.g, 学級やグループで話し合う活動を授業などで行っていましたか、などの 15 項目に 5 件法で回答）、小学校のほうがかなり積極的に取り組んでいるのに対し、中学校ではばらつきがある（図 8：

アクティブラーニングの実施（戸田市）。

図 6：学校の運営体制（戸田市）

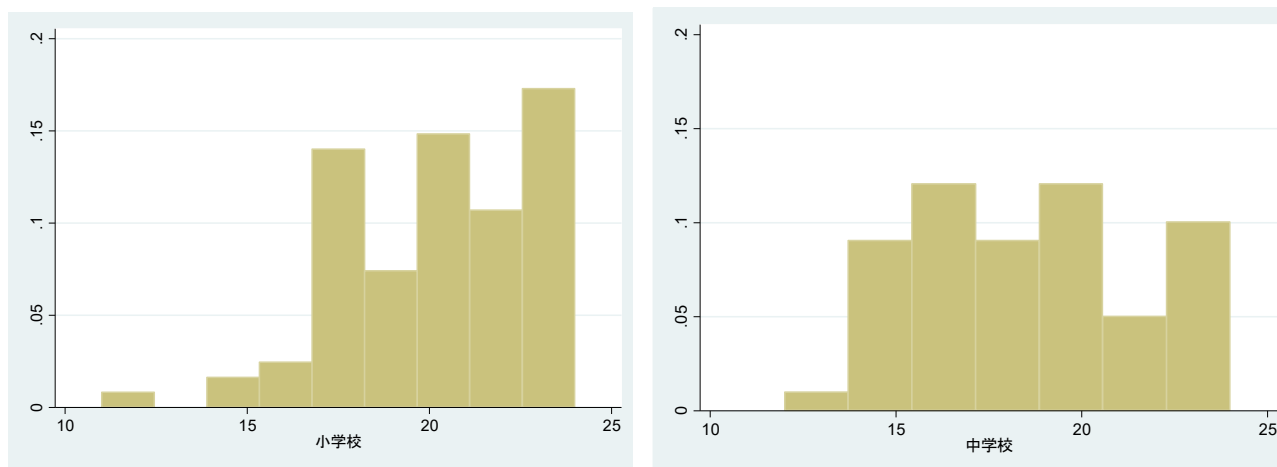


図 7：教員の協働体制（戸田市）

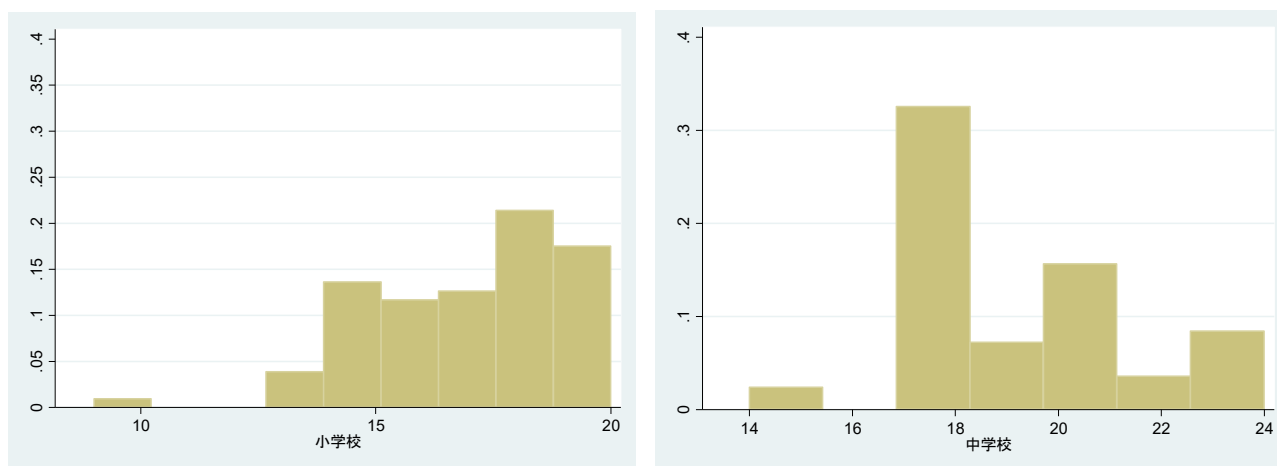
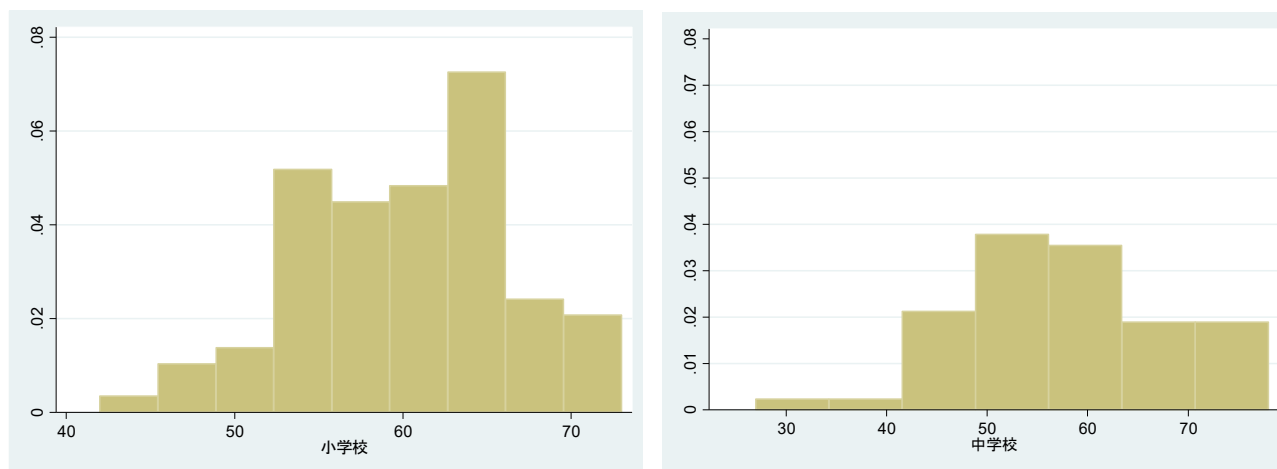


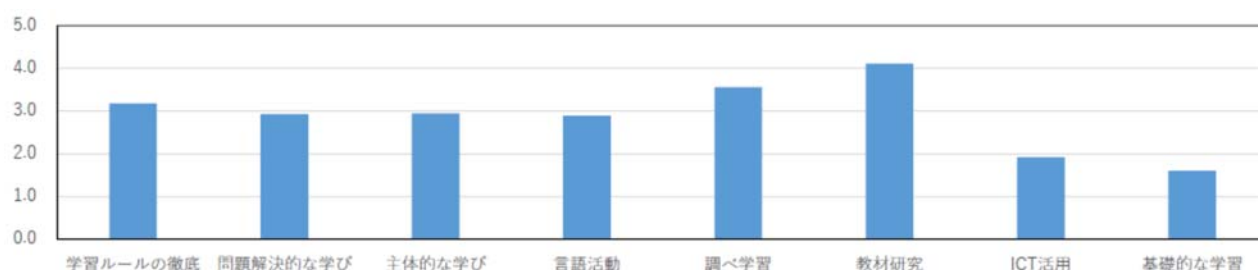
図 8：アクティブラーニングの実施（戸田市）



どのように力点をおいて指導しているかという質問に対しては、図 9 のとおりアクテ

ィブラーニングと考えられる問題解決的あるいは主体的な学びよりも、各科目の教材研究を行ってそれを指導に活かしていると回答している教員が多い。

図 9：教員が力を入れている指導（戸田市）



一方、図 9 の教員が力点を置いている指導の方法と、生徒の学習方略との相関関係をみると、表 17：学習方略と指導の関係（戸田市）で明らかとなっており、教材研究は学習方略に影響しておらず、むしろ問題解決的な指導によって学習方略（特に、プランニング方略、作業方略、努力方略）が変化していることがわかる。

表 17：学習方略と指導の関係（戸田市）

	ルール	問題解決	主体的	言語活動	調べ学習	教材研究	ICT活用	基礎的
学習方略	－	△	－	－	－	－	－	－
柔軟的方略	－	－	－	△	－	－	－	－
プランニング方略	－	△	－	－	－	－	－	－
作業方略	－	△	－	▼	－	－	－	－
人的リソース方略	－	－	－	－	－	－	－	－
認知方略	－	－	－	－	－	－	－	－
努力方略	－	△	－	－	－	－	－	－

（注）推計式では生徒の学年と性別を制御している。図中の△は正で統計的に有意、▼は負で統計的に有意、－は統計的には有意ではない関係を示す。
（出所）埼玉県戸田市で実施した教員調査（第1回）

戸田市の調査においても、学習方略は学力に影響しており(表 18)、教員の指導法は学習方略の変化させることを通じて、学力に影響する可能性がある。

表 18：学習方略と学力の関係（戸田市）

	国語	算数・数学
柔軟的方略 ＝学習の進め方を自分の状態に合わせて柔軟に変更していく方略	0.001 (0.006)	0.005 (0.005)
プランニング方略 ＝計画的に学習に取り組もうとする方略	0.014* (0.006)	0.003 (0.005)
作業方略 ＝（ノートに書いたり、声に出したり）作業を中心として学習を進める方略	-0.009 (0.006)	-0.007 (0.005)
人的リソース方略 ＝友人を利用して学習を進める方略	-0.021*** (0.005)	-0.023*** (0.004)
認知的方略 ＝理解や精緻化、集中力と言った認知的な働きを重視して学習を進める方略	0.020** (0.007)	0.018** (0.006)
努力調整方略 ＝「苦手」などの感情をコントロールして学習への動機を高める方略	0.031*** (0.006)	0.044*** (0.005)
N	2,975	2,901

(注) 1. この回帰分析では、平成27年度の能力パラメタ、生徒の性別（女性0、男性1）、学年、生まれ月、保護者の社会階層（＝生活保護・就学援助受給世帯であれば0、その他世帯であれば1）、担任教員の性別（女性0、男性1）、担任教員の勤続年数を制御している。
2. 標準誤差は分散不均一を修正している。
3. ***は0.1%水準で、**は1%水準で、*は5%水準で統計的に有意であることを示す。
(出所) 埼玉県学力・学習状況調査（戸田市）

本稿の分析結果を解釈する上での留意点として、本稿の回帰分析から得られた結果は全て「相関関係」に過ぎず、「因果関係」を示唆するものではないということである。2年間にわたって行われた調査ではあるが、非認知能力や学習方略などの質問は2016年度から開始されたことから、観察不可能な異質性を制御できるようなモデルでの教育生産関数の推定は、現時点では困難である。来年度以降、データが蓄積されればより信頼性の高い分析が可能になろうし、今年行われた様々な分析を踏まえて、学力や非認知能力を上げる因果効果が期待できそうな介入（例えば、教員研修や授業実践）を実施して、その効果測定を行えば、より厳密な評価が可能になろう。

参考文献

- Ammermuellepr, A. (2005), Poor Background or Low Returns? Why Immigrant Students in Germany Perform so Poorly in PISA, ZEW Discussion Paper No. 05-18, Mannheim.
- Aslam, M., & Kingdon, G. (2011). What can teachers do to raise pupil achievement?. *Economics of Education Review*, 30(3), 559-574.
- Bietenbeck, J. (2014). Teaching practices and cognitive skills. *Labour Economics*, 30, 143-153.
- Coleman, J. S., & Department of Health USA. (1966). *Equality of educational opportunity* (Vol. 2). Washington, DC: US Department of Health, Education, and Welfare, Office of Education.
- Cunha, F., Heckman, J. J. and Schennach, S. M. (2010). Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica*, 78(3), 883-931.
- Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92, 1087-1101.
- Embretson S. E., and Reise, S. P. (2000). *Item Response Theory for Psychologists*, Mahwah, NJ : Psychology Press.
- Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The effect of school resources on student achievement. *Review of educational research*, 66(3), 361-396.
- Hanushek, E. A., Rivkin, S. G., & Taylor, L. L. (1996). Aggregation and the Estimated Effects of School Resources. *The Review of Economics and Statistics*, 611-627.
- Hanushek, E. A. (1989). The impact of differential expenditures on school performance. *Educational researcher*, 18(4), 45-62.
- Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational evaluation and policy analysis*, 19(2), 141-164.
- Hanushek, E. A. (1994). Money might matter somewhere: A response to Hedges, Laine, and Greenwald. *Educational Researcher*, 23(4), 5-8.
- Heckman, J. J., & Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the GED testing program. *The American Economic Review*, 91(2), 145-149.
- Heckman, J. J., & Krueger, A. B. (2005). Inequality in America: What Role for Human Capital Policies?. *MIT Press Books*, 1.
- Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor economics*, 24(3), 411-482.
- Heckman, J., Moon, S. H., Pinto, R., Savelyev, P., & Yavitz, A. (2010). Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program. *Quantitative economics*, 1(1), 1-46.
- Hedges, L. V., Laine, R. D., & Greenwald, R. (1994). An exchange: Part I*: Does money matter? A meta-

- analysis of studies of the effects of differential school inputs on student outcomes. *Educational researcher*, 23(3), 5-14.
- Kawaguchi, D. (2016). Fewer school days, more inequality. *Journal of the Japanese and International Economies*, 39, 35-52.
- Matsuoka, R., Nakamuro, M., & Inui, T. (2015). Emerging inequality in effort: A longitudinal investigation of parental involvement and early elementary school-aged children's learning time in Japan. *Social science research*, 54, 159-176.
- Oshio, T. & Senoh, W. (2007). The economics of education in Japan: A survey of empirical studies and unresolved issues, *Japanese Economy*, 34(1), 46-81.
- Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of educational psychology*, 82(1), 33.
- Pritchett, L., & Filmer, D. (1999). What education production functions really show: a positive theory of education expenditures. *Economics of Education review*, 18(2), 223-239.
- Tsukayama, E., Duckworth, A. L., & Kim, B. (2013). Domain-specific impulsivity in school-age children. *Developmental Science*, 16, 879-893.
- von Davier, M., Gonzalez, E., & Mislevy, R. (2009). What are plausible values and why are they useful. *IERI monograph series*, 2, 9-36.
- Woessmann, L. (2008). How equal are educational opportunities? Family background and student achievement in Europe and the United States. *Zeitschrift Für Betriebswirtschaft*, 78(1)(1284), 45rtsc
- 赤林英夫・直井道生・敷島千鶴(2016)「学力・心理・家庭環境の経済分析-全国小学生の追跡調査から見えてきたもの-」,有斐閣.
- 加藤健太郎・山田剛史・川端一光(2014)「Rによる項目反応理論」, オーム社.
- 北村行伸 (2006)「パネルデータの意義とその活用—なぜパネルデータが必要となったのか」『日本労働研究雑誌』 48(6), 6-16
- 北村行伸 (2013)「パネルデータの分析手法の展望」『季刊家計経済研究』 No.100, pp.60-69.
- 国立教育政策研究所 (2016)「平成 28 年度 全国学力・学習状況調査 調査結果資料」
<http://www.nier.go.jp/16chousakekkahoukoku/factsheet/16prefecture/11_saitama/index.html> (2017 年 2 月 19 日アクセス)
- 国立大学法人お茶の水大学編 (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』.
- 埼玉県 (2016)「埼玉教育委員会」< <https://www.pref.saitama.lg.jp/kyoiku/>> (2017 年 2 月 19 日アクセス)
- 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—全国学力・学習状況調査(きめ細かい調査)の結果を活用した実証分析—」,『国立教育政策研究所紀要』第

145 集, p.119-129.

東京都教育委員会 (2016)「平成 28 年度「児童・生徒の学力向上を図るための調査」を実施しました」<<http://www.kyoiku.metro.tokyo.jp/buka/shidou/chosa/pr160707.html>>(2017 年 2 月 19 日アクセス)

北條雅一 (2011)「学力の決定要因—経済学の視点から」『日本労働研究雑誌』No. 614, p16-27

星野崇宏・岡田謙介 (2016)「欠測データの統計科学:医学と社会科学への応用」岩波書店

前馬優策 (2016)「授業改革は学力格差を縮小したか」志水宏吉・高田一宏 編「マインド・ザ・ギャップ! —現代日本の学力格差とその克服—」(pp. 81-106), 大阪大学出版会.

村木英治(2011)「項目反応理論」, 朝倉書店.

文部科学省 (2011)「平成23年度以降の全国的な学力調査の在り方に関する検討のまとめ」
<http://www.mext.go.jp/b_menu/shingi/chousa/shotou/074/toushin/1304351.htm>
(2017年2月19日アクセス)

文部科学省 (2013)「平成25年度実施の都道府県・指定都市に寄る独自の学力調査について」
<http://www.mext.go.jp/a_menu/shotou/gakuryoku-chousa/sonota/detail/1335150.htm>
>(2017年2月17日アクセス)

文部科学省 (2016)「全国学力・学習状況調査小兵データの貸与の在り方について(案)」
<http://www.mext.go.jp/b_menu/shingi/chousa/shotou/112/shiryo/attach/1377298.htm>(2017年2月19日アクセス)

須藤康介 (2013)「学校の教育効果と階層」, 東洋館出版.