

学力調査における認知能力の捉え直し

調査方法・社会的経済的地位・非認知能力から

田端 健人

The Redefinition of Cognitive Skills in Academic Achievement Researches

Focusing on the Methodology, Socio-Economic Status and Noncognitive Skills

TABATA Taketo

概要

本稿は、「学力」を「テストで測定される能力」と定義した上で、学力調査分析の近年の動向をレビューする。まず、古典的テスト理論（CTT）の問題点を指摘し、項目反応理論（IRT）やパネルデータの活用など近年の学力調査の動向と特徴を概観する。こうした調査理論のメタ分析から、「観測する理論が異なれば、観測される事実が異なる」という独自の哲学的知見が得られる。次に、学力調査の主要な分析結果のうち、児童生徒の学力と保護者の社会経済的地位（SES）との相関関係に焦点を絞る。従来の研究ではSES-学力規定説が有力であったが、近年の学力調査分析ではその限界が実証されている。とりわけ、日本子ども調査パネルの分析は、クロスセクションデータで確認される学力-SES相関が、観測不可能な固定的要因を制御すると消失するという驚くべき結果を示している。最後に、近年注目される「非認知能力」について概観し、この要因を学力調査に加味することでSES-学力規定説がいつそう限定されることを示す。以上の学力調査分析のメタ分析から、「学力に関して、測定観測の理論と方法が変化することにより、測定観測される事実が存在レベルで変化する」という「観測者問題」が結論づけられる。

Key words: 項目反応理論 SES（社会経済的地位） 観察不可能な要因 観測者問題

はじめに

「学力」とは何か。この根源的な問いは、「テストで測定される能力だけが学力か」という反語的な問いとしばしば連動している。こうした難問を徒手空拳で議論しても、得るところは少ないだろう。割り切って、「学力とは、学力テストで測定される能力である。」こう定義することから始めよう。

学力をこのように定義すると、「学力とは何か」という大問題を一時的に棚上げできる。そうすると「学力」をめぐる、次のような問いが噴出する。テストの得点が高い児童生徒と、そうでない児童生徒との差は、何によって決まるのか（学力の決定要因は何か）。学力の高低は、児童生徒の家庭環境で決まるのか、それとも家庭環境の格差を乗り越え、学力を向上させる効果的な学校や教師は存在するのか。存在するとすれば、それはどのような学校や教師なのか。

学力調査は、こうした問いに答えるために実施されている。個人と集団の学力をただ測定したいという単純な測定願望で測定しているわけではない（測定のための測定ではない）。競争をあおるためでもない（結果的にそうになっていたとしても）。学力調査は、現在の学力を過去の学力と比較し、それが向上したか維持されたか低下したか（学力の動態）を把握するためであり、その変動の決定要因を把握するためであり、そこか

ら将来の効果的な学力向上を図るためである。それゆえ学力調査は、これらの目的を達成するために適切に設計され実施されなければならない。

I 学力調査の設計をめぐる議論

近年、学力調査の設計そのものを問い直す議論が活発化している。その一つが、テストや統計を専門とする研究者からの「全国学力・学習状況調査」（以下「全国学力調査」）への批判である。2007（平成19）年以降、毎年約60億円の国費を投じて実施される全国学力調査の、いったい何が問題なのか。

1 全国学力・学習状況調査の問題点

最大の問題の一つは、この学力テストでは、学力の経年比較ができないことである（cf., 川口, 2014, p. 2）。これは正直、非常に驚くべき問題の指摘である。これが本当ならば、異なる集団、例えば、2019年度の6年生のテスト結果と、2018（平成30）年度の6年生のテスト結果を比較することもできないし、同一集団、例えば2016年度の小学6年生と2019年度の中学3年生を比較することもできない。そのため、このテスト結果からは、前の年から子どもの学力が向上したとも低下したともいえず、ましてや学力の向上や低下にどのような施策や指導が効果的かを特定することもできない。世間では、「〇〇自治体では、△△に取り組むことで、平均正答率が全国平均との差でプラス××に向上した」といった言説がしばしばなされるが、全国学力調査から単純にこのように断言することは間違っている。同様に、「X小学校が、全国学力調査でこれまで全国平均以下から全国トップクラスになった」としても、残念ながら、X小学校の学力が向上したかどうかは、実はわからない¹。

もう一つの問題は、質問紙調査の学習状況調査が不十分であり、特に児童生徒の家庭環境がほとんど把握できないことである（cf., 川口, 2014, p. 2）。これまでの研究で明らかにされてきたのは、教育現場での教師たちの実感とも符合するように、学力と家庭環境には正の有意な相関がある、ということである。にもかかわらず全国学力調査には、家庭環境を正確に知るための保護者への質問紙調査がなく、その相関を正確には把握できない²。

私たちが素朴に「家庭環境」とか「家庭的背景」と呼ぶ、保護者の収入や学歴や教育熱などは、学術的には、「社会経済的地位（Socio Economic Status）」、略して「SES」と表現される。本稿でも以下キーワードになる。

子どもの学力の決定要因は、その保護者のSES、学習意欲、教師の働きかけ、友達の影響など、多岐にわたる。こうした決定要因の全体をバランスよく視野に収めるためには、経済学で標準的に用いられる「教育生産関数」モデルが有効である。教育生産関数は次のように定義される（以下表1まで、石川ほか, 2017, p. 4より引用）。

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

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 は時間を表している。

¹ 文部科学省や国立教育政策研究所がこの問題に無自覚だというわけではない。文部科学省は、後述のIRTモデルによる経年変化分析を、抽出法により2013（平成25）年度に導入し、2016（平成28）年度に2回目を実施している（cf., 文部科学省ほか, 2017, p. 2; 東北大学, 2018, i）。

² こちらに関しても文部科学省は、全国学力調査の追加分析調査として、2013（平成25）年と2017（平成29）年に無作為抽出法で「保護者に対する調査」を実施した。平成29年度調査概要については、国立教育政策研究所のウェブサイトに掲載されている。

(https://www.nier.go.jp/17chousakekkahoukoku/kannren_chousa/hogosya_chousa.html)

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 (社会的経済的地位)	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)

教育政策や指導を効果的にする議論のためには、学力向上の阻害要因や促進要因を特定する必要がある、そのためには、学力テストの結果だけでなく、上記の多岐にわたる投入要素に関する情報が不可欠である。それゆえ、児童生徒や保護者や教師へのアンケート調査をセットにして実施する必要がある。ところが、日本では最近になるまで教育生産関数の推定がほとんどおこなわれてこなかった (cf., 石川ほか, 2017, p. 5)。全国学力調査しかりである。

上記第一の問題、すなわち学力の経年変化が把握できないという大問題は、テストを設計する理論上の問題に起因する。学力の理論と実践を往還させる一つの要になるため、現職教員や教育学部学生にもその初歩の初歩を知っておいてほしい。

2 古典的テスト理論 (CTT)

全国学力調査は、日本でのこれまでのテストの主流であった古典的テスト理論 (Classical Test Theory: 以下 CTT) で設計されている。以下主に、文部科学省ほか (2017, pp. 27-29) と石川ほか (2017, pp. 8-9) をもとに、CTT の性格を概観しておこう。

CTT では、テスト結果 (観測値: x) は、受検者 (児童生徒) の能力を反映した真の値 (真値: t) と、受検者の当て推量や採点のブレなどランダムな誤差 (e) とから構成される、と仮定している。数式で表現するなら、

$$x = t + e$$

と表せる。

CTT は、誤差 (e) をできるだけゼロに近づけ、素点 (x) が真値 (t) の代理変数となることをめざしているが、これはそれほど簡単なことではない。誤差の影響を緩和する方法の一つに、同じテストを同一集団で複数回実施する「再テスト法」という信頼性係数の推定値を求める方法があるが、同じテストを複数回受けることによる練習効果によって、測定されるべき真値そのものが変化する問題がある。また、ある受検者はテスト問題をはじめから解き、時間が足りなくなって、後に解いた問題項目を当てずっぽうに解答した場合には、後のテスト項目の真値が測定できない。問題項目の出題順序の影響が、CTT では考慮されていない³。

さらに致命的なことに、CTT では出題範囲が制限されるという欠点を克服できない。これに対し、以下で紹介する IRT は、少数の問題に様々な範囲や難易度の問題を組み込んだ問題冊子を利用できる⁴。

最大の問題は、CTT テストは、国語や算数・数学などの同じ教科の同じ事項を扱う問題であっても、テスト問題が違えば、異なるテスト間の比較ができないことである。CTT では、「問題が易しかったから正解が多か

³ 出題順序の影響を除外し信頼性を確保する方法には、折半法、またいっそう確かな手法として、クロンバックの α 係数 (信頼性係数) の活用がある。ただ、 α 係数も万能ではなく、入試のような高度の精度が求められる場面では問題が残るといえる。これは査読者からの示唆である。

⁴ これも査読者からの貴重な教示である。この場をお借りして感謝申し上げたい。

ったのか、受検者の能力が高いから正解が多かったのか」を「判断できない」（光永, 2018, p.98）。異なるテスト間の難易度を同じにする「等化 (equating)」が、CTT では難問なのである。

「素点を標準化したり、偏差値へ変換することで、複数のテスト間の比較が可能になるだろう」と反論したくなるかもしれないが、CTT の標準化には無視できない限界があるという。標準化は、複数のテストの得点分布を、平均値と標準偏差（ばらつき）の二つの指標のみを使って比較可能にすることをめざす。しかし、実際のテストの分布は、平均値を中心の山とする左右対称（正規分布）ではなく、歪んでいたり、山が二つであったりする。それゆえ、本来ならば、歪みを表す指標や、山の個数を表す指標など、様々な指標を考慮しなければ、分布を揃えることができない。加えて「標準得点は、分布の中での相対位置を表すのには有効だが、分布の平均値をそろえることによって分布全体の横軸方向の移動の影響を消去しているため、学力がどう変化したかを比較するには適していない。また、分布の元となる集団が異なれば標準得点も変わってくるため、あくまで同一集団（と仮定される）の場合にのみ比較可能な指標であるという限界がある。」（文部科学省ほか, 2017, p.28）

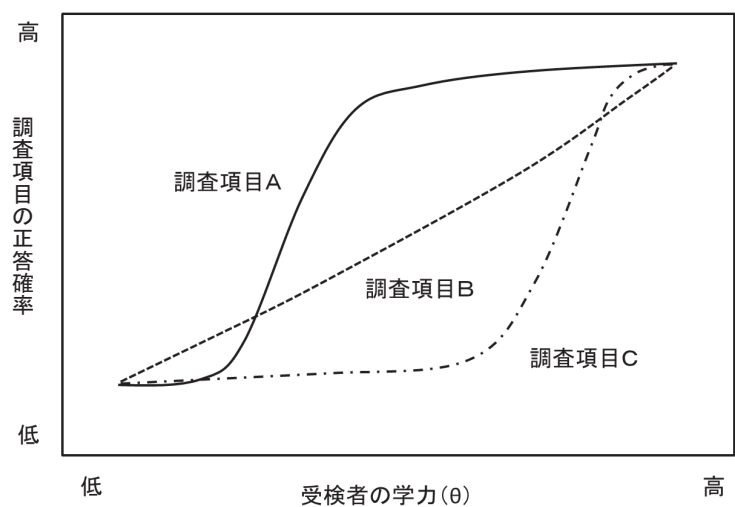
CTT のこうした問題を解決するのが、項目反応理論（Item Response Theory : 以下 IRT）である。国際学力調査 PISA や TIMSS、英語能力試験 TOEFL などは、IRT をもとに設計されている。そのため、異なるテスト、異なる集団の間でも、特定の能力の比較が可能になるという。

3 項目反応理論（IRT）

CTT では、受検者の学力は、真値という仮定から、テスト得点で表される。ところがテスト得点は、テスト問題（項目）の難易度に左右される。しかし CTT はテストの難易度と受検者の学力とを切り離すことができない。いわば CTT の「テスト得点」は、学力の真値を表現したいと願いながらも、真値ではなく、真値とテスト問題の難易度との混交を表現している。しかも「テスト問題の難易度」は、CTT では、正答率によって表される（例えば、同一集団で正答率 30%の問題は正答率 90%の問題よりも難しい）。ところが、正答率は、受検者集団を分母としているので、CTT テストの難易度は、原理的に受検者集団に依存する。同じ問題であっても、学力の高い集団なら、例えば正答率 90%の易しい問題になるが、学力の低い集団なら、正答率 30%の難しい問題になる。要するに CTT のテスト得点は、学力の真値とテスト問題の難易度とを区別できず、難易度は受検者集団に依存する。

これに対し、「IRT モデルには、受検者の学力を表す受検者母数とむずかしさなどの項目の特性を表す項目母数が含まれて」おり、「テストの結果から、項目に依存しない受検者母数〔受検者の学力〕と受検者集団に依存しない項目母数〔難易度など〕を推定することができる」（東北大学, 2011, p. 15, []内引用者）。それゆえ、異なるテストを受検した受検者どうしの学力を、同一尺度上で比較できるという。

IRT は、問題項目ごとに、①項目の難易度（困難度）、②項目の識別力、③受検者の学力（尺度値 θ ）を、次のような項目特性曲線により、それぞれ切り分けて表現可能にす



(文部科学省ほか, 2017, p.30)

図 1 : 項目特性曲線 (模式図)

る（図1参照）。

問題項目 A は、曲線の立ち上がりの位置が学力（ θ ）の低い側にあり、高学力の受検者の正答確率が高い、つまり簡単な問題であることを示している。対して調査項目 C は、難しい問題である。この難易度は受検者集団には依存しない。また調査項目 A と B を比較すると、A は B よりも曲線の傾きが急である。この傾きは問題の「識別力」を表現している。A は、ある学力以下の受検者の正答率は低いが、ある学力以上になると正答率が高いため、特定の学力が有るか無いかその閾値を鋭敏に識別する。対照的に項目 B は、閾値の識別力がない。

項目特性曲線は、いくつかの数式に変換可能であるが、2パラメタ・ロジスティック・モデルと呼ばれる以下の数式がその一つとされる。 i 番目の受検者の学力（尺度値）は θ_i 、 k 番目の調査項目の識別力は a_k 、その項目の困難度は b_k で表される。⁵

$$P(x_{ik} = 1|\theta) = \frac{1}{1 + \exp(-1.7a_k(\theta_i - b_k))}$$

- x_{ik} は、 i 番目の受検者が k 番目の調査項目に正答できるかどうかを表し、正当を 1、誤答を 0 とする。
- $P(x_{ik} = 1|\theta)$ は尺度値が θ_i である受検者が、 x_{ik} に正答する条件付き確率。
- 識別力 a_k の値が大きくなると、曲線の傾きが急になり、識別力が高くなる。

(cf., 文部科学省ほか, 2017, pp.30-31)

調査項目の特徴である識別力 a_k と難易度 b_k が未知の場合は、「最尤（さいゆう）法」という数学的手法で推定する。これら 2 つのパラメタがわかれば、それらの組み合わせと受検者の解答状況から、当該受検者の学力（尺度値） θ_i が推定可能になる（cf., 文部科学省ほか, 2017, p.31）。また、IRT で設計された異なるテストならば、項目の一部あるいは受検者の一部を共通にすることで、異なるテストの関係性が算出可能になり、複雑な数式による変換によってテストの「等化」ができる（cf., 文部科学省ほか, 2017, pp.32-34）。これにより、異なる時点でおこなわれたテスト間の比較、複数の受検者・回答者集団間や異なる年齢間での学力の比較が可能になる。⁶

4 クロスセクションデータとパネルデータ

国際学力調査 PISA や TIMSS は、特定の年齢の児童生徒を対象に、年度を超えて実施される。つまり調査される集団は常に異なる。このように異なる集団を対象とする調査は、「クロスセクションデータ」と呼ばれる。クロスセクションデータを比較可能にして、経年変化を把握する一つのテスト理論が IRT である。

他にも、経年変化を捉える方法がある。同一の子どもや同一集団（学校や自治体）の変化を追跡する「パネルデータ」という調査方法である。北村行伸によれば、アメリカの経済学領域ではパネルデータを利用した論文が急速に増加しているという（cf., 北村, 2013, pp.60-61）。パネルデータの利点は幾つかあるが、その一つは、観察点が格段に増加することによる推定精度の向上である（cf., 北村, 2013, p.61）。次に、「観察不可能な変数」を「固定効果として抽出」できることである（cf., 北村, 2013, p.61）。学力調査を例にすれば、個人の生まれつきの能力とか、学校・学級経営の質とか、地域の教育風土とかの要素（変数）は、観察が不可能ではないにしても、学力テストのような数値化がはるかに難しい。計量経済学ではこれを

⁵ 2PLM の係数は、正確には 1.702 とすべきかもしれないが、本稿では、初心者向きに開設した文部科学省, 2017 (p.33) や光永, 2018 (p.127) に従った。

⁶ 査読者の示唆によれば、このように難解な IRT でなくても、旧来の S-P 表を利用すれば、IRT とほぼ同じ分析ができるという。これまでの学力テストのデータの蓄積を生かすという点から、S-P 表の活用は非常に意義深い。

「観察（観測）不可能な変数」と呼ぶようだが、パネルデータでは、この変数を固定効果として捉えることで、抽出できるというのである。さらにパネルデータは、マクロデータとは異なり、「研究者が全ての個票の数値を観察できるので推計上の問題に関してもさまざまな解決方法を考えることができる」（北村，2013，p.62）という利点をもつ。この点は、クロスセクションデータでも、PISA や TIMSS のように、個票を広く公開すれば可能になる。その重要性に鑑み、全国学力調査でも、2018 年より研究目的で個票が貸与されるようになった。

5 代表的な学力調査の一覧

テストの設計に関わる以上の観点から、日本で実施されている代表的な学力テストの一覧を示しておきたい。この一覧を見れば、2015（平成 27）年から開始されている埼玉県（さいたま市を除く）の学力・学習調査の先進性も理解できる。

表 2：日本で実施されている代表的な学力調査

	全国学力調査	PISA/TIMSS	埼玉県学力調査
測定の方法	CTT	IRT	IRT
調査方法	悉皆	抽出	悉皆
対象学年	小 6、中 3	PISA：15 歳 TIMSS：小 4、中 2	小 4～中 3
対象者数	約 217 万人 (2019 年度)	PISA(2018)：約 60 万人 (うち日本：約 6,100 人) TIMSS(2015)：約 52 万人 (うち日本：9,100 人)	約 29 万人
調査サイクル	毎年	PISA：3 年ごと TIMSS：4 年ごと	毎年
データ	クロスセクション	クロスセクション	パネル
テスト項目の公開／非公開	公開	一部非公開	非公開
個票の公開／非公開	2018（平成 30）年度より研究者に貸与	公開	研究者に公開
質問紙調査	児童生徒 学校	児童生徒	児童生徒 教育委員会・学校

（石川ほか，2017，p.20 と川口，2018，p.1 を参照し田端が作成）

埼玉県学力・学習調査（以下「埼玉県学力調査」）では、児童生徒と各教育委員会と学校を対象に質問紙調査を合わせて実施している。生徒対象の質問紙調査には、「非認知能力」「学習方略」「教員の指導」「保護者の社会的経済的地位」を問う項目がある⁷。先の教育生産関数の投入要素がバランスよく含まれている。

非認知能力と学習方略を問うのは画期的に見える。この調査では、非認知能力として、「①自制心」「②自己効力感」「③勤勉性」「④やり抜く力」とカテゴリー化される項目がある（埼玉県，2019，第 2 章）⁸。学習

⁷ 非認知能力の項目を加えたのは、2016（平成 28）年度からである。

⁸ 埼玉県学力・学習状況調査は、2015（平成 27）年度から、現時点で最新の 2019（令和元）年度までの全て

方略としては、「①柔軟の方略…自分の状況に合わせて学習方法を柔軟に変更していく活動」「②プランニング方略…計画的に学習に取り組む活動」「③作業方略…ノートに書く、声に出すといった、『作業』を中心に学習を進める活動」「④人的リソース方略…友人を利用して学習を進める活動」「⑤認知的方略…より自分の理解度を深めるような学習活動」「⑥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意欲を高める活動」など、多彩な項目が盛り込まれている。

ただ、保護者の SES については、保護者本人に尋ねていないことや、職業や年収などを直接質問していないこと、代理変数（代替指標）として使える項目もわずかであること（家庭の蔵書数、通塾時間のみ）など、必ずしも十分ではない。また教員の質や指導法についても、学校や教員への質問でなく⁹、児童生徒の印象の代理変数しかないこと、それも項目として限られており、教師の授業方法、学級経営、学校マネジメントなどの具体がわかる項目がないことなど、十分とはいえない。

とはいえ、埼玉県の問題紙調査は、全国学力調査と比較して格段に豊富な項目になっている。また埼玉県教育委員会は、この調査で「大きな伸びが見られた学級や教科の担当者からの聞き取りや授業参観」もおこない、報告書で取り組みを紹介している（cf., 埼玉県, 2019, はじめに; 第4章）。

II 「学力」と「社会的経済的地位（SES）」

1 埼玉県学力・学習状況調査から

埼玉県学力調査に関する石川ほか（2017）と埼玉県（2019）の分析結果からは、多くの知見や示唆がえられる。なかでも学校関係者にとってショックなのは、保護者の SES と学力格差との相関である。石川善樹たちの分析では、SES は、児童生徒に対する次の二つの質問項目を代理変数としている（cf., 石川ほか, 2017, p. 36）¹⁰。

- ・ 蔵書数を尋ねる項目「家には、自分や家の人が読む本がどれくらいありますか」。回答は次の5つから選択。「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）」。
- ・ 通塾を尋ねる項目「学習塾（家庭教師に教わっている場合もはいます）で1週間のうち、どのくらいの時間、勉強しますか」。回答は次の8つから選択。「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」。

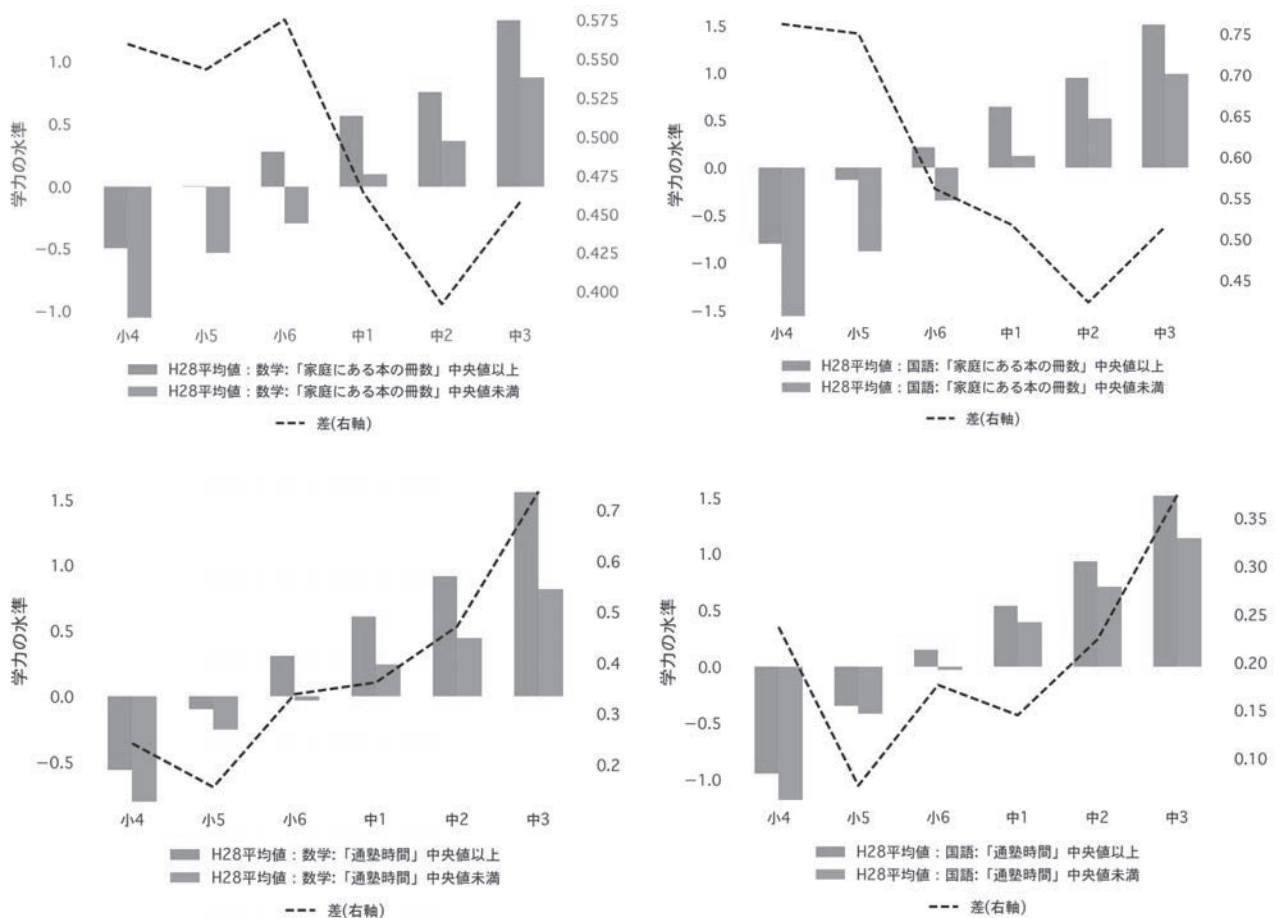
この2つの変数の中央値をとり、中央値よりも高いグループを SES 高グループ、中央値よりも低いグループを SES 低グループと定義し、2つのグループ間で学力を比較したのが、次の図2である（cf., 石川ほか, 2017, p. 37）。上が蔵書数、下が通塾時間、左が数学、右が国語の学力との相関である。

の報告書が、ホームページ上で公開されている。ただその報告書にはページ数が記されていないため、引用は、掲載されている「章」を明示した。

⁹ 学校と市町村教育委員会への調査項目の具体は、ホームページ上では確認できなかった。

¹⁰ ただし、通塾の有無や時間が、SES をどこまで反映しているかは議論の余地がある。例えば、全国学力調査をもとにした2016年度の中学生の通塾率ランキングで、埼玉県は8位（67.3%）であり、秋田県は45位（30.1%）とされる。同等の経済状況でも、通塾の有無や長短は、家庭の経済状況だけでなく、それぞれの地域風土や私立学校の有無などに依存するため、通塾時間がどこまでSESの代理変数となるかは定かではない。SESについては、次節で紹介する「日本子どもパネル調査」のように、保護者に直接尋ねる方が格段に精度が高い。なお、通塾率ランキングについては、以下の「Gaccomm 学校教育情報サイト」を参照した。

（https://www.gaccomm.jp/?m=pc&a=page_o_region_ranking&info_kind=juku_go_ratio_js&lang=ja&ssl_param=1）



(石川ほか, 2017, p. 37, 図表番号のみ引用者が変更)

図 2 : 保護者の社会的経済的地位による学力格差

この図にも明らかなように、SES 高グループ (青色の棒グラフ) と SES 低グループ (灰色の棒グラフ) とでは明確な学力水準の差があり、「この差は統計的に有意である」(石川ほか, 2017, p. 37)。また、SES と関連するこの学力差は、既に小学 4 年生の時点で顕在化している (cf., 石川ほか, 2017, p. 37)。またどの学年でも、いずれの教科でも、SES が低いグループが SES の高いグループよりも学力が高いことはない。これは、学校教育が親の SES のハンディキャップを超えて、子どもの学力格差を解消するほどには機能していない、という失望感を強くする結果である。

さらに図 2 からわかるように、下段の通塾時間については、2 つのグループの差 (波線とグラフの右側のポイントで表される) は、学年が上がるにつれ高まっており、「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」(cf., 石川ほか, 2017, 38)。

しかし石川たちの指摘にはないが、上段の蔵書数のデータでは、SES 高グループと低グループとの差を表す波線は、全体として学年進行とともに減少傾向にある。これは学年が上がるにつれ、蔵書数が異なる 2 つのグループの学力差が縮小していることを意味する。蔵書数を SES の代理変数と認めるならば、学年が上がるにつれ、学力差に対する SES の影響が小さくなっていると推論できる。SES 低グループは通塾していなかったり、その時間も短いので、学力形成に影響を与える要素は、学校と教師が優勢である。それゆえ上段のデータからすれば、学校は、SES 低グループの学力差を下支えし、SES 高グループとの学力格差を縮小するの

に貢献しているとも推測できる。通塾時間の格差が学年を追うごとに増加していることを加味すれば、また上段と下段のグループが同じであると仮定すれば、次のようなストーリーを描くことも不可能ではない。すなわち、SES 高グループの児童生徒は、学年が上がるにつれ、学習塾でいっそう長時間学習することで学力を上げている。SES 低グループの学習塾での学習時間との差は、学年を追うごとに開くばかりである。それゆえ、SES 低グループとの学力格差は学年を追うごとに拡大してよいはずだが、SES 低グループは、学習のための限られたリソースである学校と教師を活用することで、高グループとの学力差を縮小している。しかしこうしたストーリーを描くには、上記データだけでは不十分であろう。さらに綿密な調査分析が必要である。

2 お茶の水女子大学調査から

学力と SES との強い相関は、埼玉県学力調査だけでなく、これまでの国内外の調査からも指摘されてきた。国内でこの問題を「学業達成の階層間格差」として追求してきたのは、教育社会学者である。日本の教育社会学がこの問題に本格的に関わるようになったのは、耳塚寛明によれば、「2002 年あたり」からである（耳塚，2007，p. 24）。耳塚は、2002 年から 2007 年現在までの研究を概観し、次の 3 つの課題を指摘している。

第一に、「学力とそれを規定する諸要因との連関は、なお仮説的に提示されているにとどまる」（耳塚，2007，p. 24）。「相関」があるからといって、「規定する要因」（因果関係）とは断言できない点に要注意である。第二に、「家庭的背景（階層変数）に関して、子どもに対する自計式質問紙調査によって得られた、父母の学歴（大卒か否か）と文化的環境に関わる幾つかの質問項目にもっぱら依存して分析を行っている」が、それでは「データの信頼性」に問題が残り、家庭背景の「非常に限定された側面のみが観察されているに過ぎない」（耳塚，2007，p. 25）。第三に、「調査対象エリアの問題」（耳塚，2007，p. 25）がある。「地域的背景によって、学力格差の程度とその規定要因との関係は異なる可能性がある」ため、特に「私立学校の存在と地位」が大きな地域と小さな地域とを比較しなければならない（耳塚，2007，p. 25）。

こうした問題意識から、耳塚を中心としてお茶の水女子大学主催の JELS (Japan Educational Longitudinal Study) というパネルデータを利用した研究が開始される。JELS の報告書は、第 1 集 (2004) から第 18 集 (2015) まで、インターネット上に公開されている。その成果はまた、耳塚 (2013)、耳塚 (2014)、中西 (2017) など著作にもまとめられている。

JELS2003 の分析成果の一つは、学力と「家庭的背景との関連は、地域によって一様ではない」（耳塚，2007，p. 32）ことを実証した点にある。これは、対照的な二つの都市（A エリアと C エリア）のパネルデータの分析結果である。A エリアは関東地方大都市近郊中都市（人口約 25 万人）で私立中学への進学率は 14.2%、C エリアは東北地方小都市（人口約 9 万人）で、寄宿せずに進学可能な私立中学校は存在しない（cf.，耳塚，2007，pp. 25-26）。これら 2 エリアの小 3、小 6、中 3、高 3、各学年約 1,000 名を対象に、①児童生徒調査、②学力調査（国語、算数・数学）、③保護者調査（家庭的背景、しつけ、文化的環境、教育期待など）、④担任教員調査（教授方法、進路指導）、⑤地域、学校の状況に関するヒアリング調査、資料蒐集をおこなっている（cf.，お茶の水女子大学，2008，pp. 1-2）。回収率は、児童生徒対象は 90%を超えているが、保護者対象は 25.8～46.0%と低い（cf.，お茶の水女子大学，2008，p. 3）。

両エリアを分析したところ、大都市近郊に位置する A エリアでは、「家庭的背景」が「学力の分布」を説明するうえで大きな役割を果たすことがわかった（cf.，耳塚，2007，p. 33）。つまり、SES が良いほど学力が高い。A エリアでは、「家庭での学習時間（努力）」以上に、「家計所得」「教育費投資」「保護者の強い学歴期待」が、「学力」と結びついている（cf.，耳塚，2007，p. 33）。ところが、いわゆる地方小都市 C エリアでは、「家庭的背景が学力に対する決定的な影響を持つとはいえない」（耳塚，2007，p. 32）。A エリアと比較して C エリアでは、父親の「非大卒と大卒の差はわずかでしかない」し、「家庭的背景や受験塾への通塾行動は、必

ずしも学力に決定的な規定力を持っているわけではない」（耳塚，2007，p. 29）。「地域による学力形成過程の多様性が存在」（耳塚，2007，p. 29）し、SES と学力格差との相関が強い地域と、相関が有意でない地域とが存在する。SES が児童生徒の学力を強く規定するとする仮説（以下「SES-学力規定説」）は万能ではなく、限定的であることが実証されたのである。そして、SES との相関が有意でない地域を理解するには、SES-学力規定説以外の学力形成過程を捉える理論枠組みが必要になる。

またこの分析結果で注目に値するのは、A エリアの学力が「保護者の強い学歴期待」と相関することを示した点である。A エリアでは、子どもの学力を説明する家庭背景に関する変数では、「保護者学歴期待」が1位、「受験塾への通塾の有無」が第2位であり、「世帯所得」は第5位である（耳塚，2007，p. 32）。

お茶の水女子大学はさらに、2013（平成25）年度と2018（平成29）年度の全国学力調査と、抽出保護者に対する追加調査とをもとに、家庭状況等が子どもの学力とどのように関係しているかを分析している（お茶の水女子大学，2014，2018）。これらの分析は、SES 規定説の限界をいっそう裏づけ、学力形成過程を複雑かつ多様に想定しなければならないことを示した。

例えば、SES スコアとそれ以外の変数を用いた分析では、小学校でも中学校でも、また算数・数学でも国語でも、正答率に強く影響する変数は、「SES スコア」「保護者の教育期待」「学習時間」であり、学年や教科によっては、保護者の期待が SES より影響力が強いケースもある（cf.，お茶の水女子大学，2014，pp. 61-62）。ここから分析者の山田哲也は、「一般的に、家庭的な背景が学力に与える影響には、SES が直接的に正答率を規定する経路と、SES が教育期待や学習時間など他の変数を媒介として正答率に影響を与える間接的な経路の両方を想定できる」（お茶の水女子大学，2014，p. 65）としている。そして山田は、SES が学力を規定するメカニズムに関して、二つの仮説を提示している。一つは、「学歴期待と教育戦略の違いが学力に影響を与える」説であり、もう一つは、「保護者と子どもとのコミュニケーション様式の違いがあり、そのことが学力の差につながっている」という説である（お茶の水女子大学，2014，p. 69）。

これらの仮説は、SES-学力規定説内での仮説の緻密化であるが、従来の単純な SES 規定説よりも、データにいっそう即し、現実にもいっそう即しているように見える。仮説がいっそう現実に即することで、これまで議論の俎上にあがらなかった存在が際立ってくる。それは、子どもが学校で良い点をとることや、子どもが高い学歴を獲得してゆくことに価値をおかない家庭やコミュニティの存在である。学校化された社会とは異なる価値観、人生のビジョンをもち、学校型とは異なるコミュニケーションに親しむ家族やコミュニティの存在である。

お茶の水女子大学調査（2014，2018）の単純集計では、「子どもにはできるだけ高い学歴を身につけさせたい」を尋ねる項目では、「あてはまらない」と「どちらかといえば、あてはまらない」を合わせた割合は、小学校の保護者で42.6%（2014）と36.8%（2018）、中学校で44.2%（2014）と38.7%（2018）である。また「学校生活を楽しければ、良い成績をとることはこだわらない」を尋ねた項目では、「あてはまる」と「どちらかといえばあてはまる」の割合は、小学校で56.1%（2014）と53.2%（2018）、中学校で47.3%（2014）と47.5%（2018）である。学歴や勉強にさほど重きをおかない保護者が4割前後もいることになる。もしこれが保護者の本心であるなら、誰のための「学力」なのかがわからなくなってしまう。学力の規定要因の調査や、学力保障のための施策や指導の模索は、誰が望み、誰のためにやっているのだろうか。学校教育を経験して大人になった保護者たちにとって、学力や学力向上とは何なのか、また勉強や学歴を優先しない保護者の価値観や人生観とは何なのだろうか。こうした問題についてもいっそう立ち入った調査研究が必要であろう。

お茶の水女子大学調査分析の2014年と2018年とを比較してわかるのは、2018年になって「非認知スキル」という理論枠組みが登場することである。これはこの間の学力論の変化を反映しているだろう。こうした新たな理論枠組みを指標（変数）とすることで、学力やSES など他の指標（変数）との相関が観測可能になり、

その観測と分析は施策や指導といった実践に影響を与える。つまり理論が観測を変え、観測結果と分析が実践に影響を与える。理論と実践の往還の一つである¹¹。

「非認知スキル」については、次のような知見が得られている。すなわち、非認知スキルは学力と弱い相関があり、SESとはほとんど相関がない、それゆえ、SESの高低にかかわらず（家庭のSESが相対的に低い場合でも）、非認知スキルを高めることができれば、学力を一定程度押し上げる可能性がある（cf., お茶の水女子大学, 2018, p. 6）。この知見がいつそう確かになれば、家庭や学校や各教育委員会での指導にも影響を与えるであろう。

3 日本子どもパネル調査（JCPS）から

SES-学力規定説の限界については、近年のパネルデータ分析でも指摘されている。慶應義塾大学パネルセンター設計・解析センターに2010年設立された「日本子どもパネル調査（Japan Child Panel Survey：以下JCPS）」の分析結果である。JCPSの2010年から2013年までの4年分のデータにもとづく研究成果は、赤井ほか（2016）の著作として出版されている。

本稿の文脈で衝撃的なのは、クロスセクションデータによる分析とパネルデータによる分析とを比較し、前者で観測された世帯所得と学力の相関関係が、後者では確認できなかったことである（cf., 赤林ほか, 2016, pp. 20-21, p. 59, p. 62, p. 80）。分析方法が異なると（「クロスセクションデータによる回帰分析」と「パネルデータによる固定効果モデル分析」との違い）、観測される事実関係の有無が左右される。驚きの結果である。どのような事情によるのか、簡単に見ておこう。

分析には、JCPS 4年分のデータに含まれる①小1～中3までの国語と算数・数学の学力テスト結果と、②保護者を対象とした社会的経済的属性のデータを用いている。学力テストは、JCPSが独自開発したものであり、CTTにもとづいている。そのため、「学力」は「学年サンプル内での相対的位置」であり、「学力の伸び」は「その相対的位置の移動」を意味している（cf., 赤林ほか, 2016, p. 34, p. 63）。既に見たようにCTTでは、同一集団であっても異なるテスト間の比較ができないという原理的問題がある以上、この分析での「学力の伸び」がどこまで妥当かは、議論の余地があるだろう¹²。

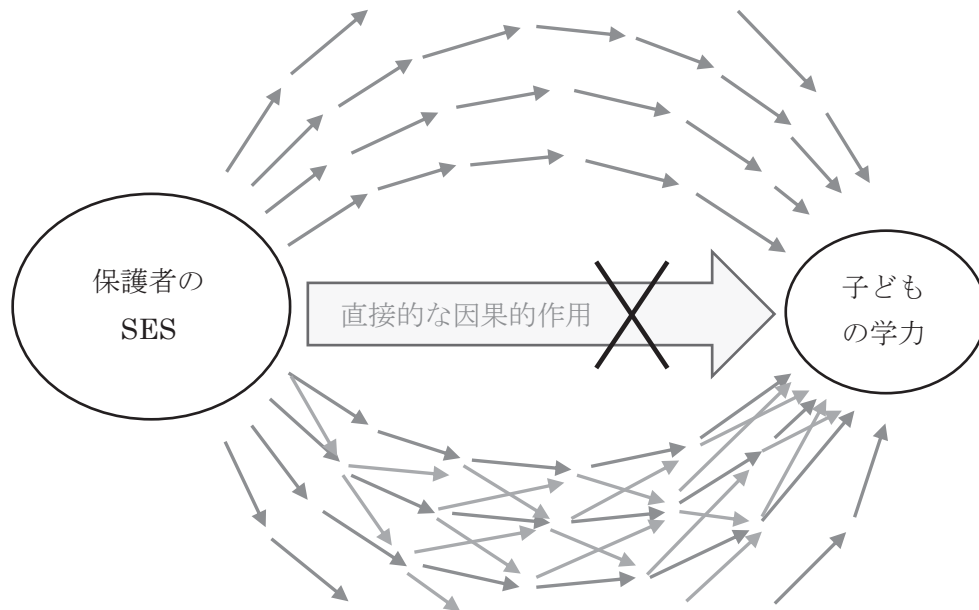
こうしたデータを利用し、学力の規定要因をクロスセクション回帰分析したところ、「世帯所得が学力に与える効果は、教科の別によらず、正かつ有意な相関を示しており、「世帯所得が学力に与える影響は、国語・数学共に学年が上がるにつれて大きくなる傾向にある」ことが示された（赤林ほか, 2016, p. 75）。これは、SES-学力規定説を支持してきたこれまでの分析結果と整合的である。

ところが、分析者の中村亮介たちは、この分析によって「世帯所得が因果的に学力に影響を与えているということとはできない」（赤林ほか, 2016, pp. 76-77）という。なぜなら、「観測できない家計や個人の属性があつて、学力と所得の両方に影響を与えている要素の存在がありうるから」（赤林ほか, 2016, p. 77）である。先に見た「観察（観測）不可能な変数」の存在である。所得と学力との相関関係は、「所得が高い家庭の子ども『だから』学力が高い」という因果関係を否定はしないが、必ずしも肯定するわけではない。両者には実

¹¹ 第I章で検討したテスト理論の違いが分析結果と知見を左右し、実践に影響するのと同様である。

¹² CTTの全国学力調査をもとに学力を経年比較した浜野隆も、次のように留保している。「平成25年度調査と29年度調査では母集団も調査問題も異なるため、単純に標準化回帰係数の数値を比較してSESの影響を論じることはできない。ただ、母集団・調査問題が異なるとはいえ、同じ国の同じ学年の全国調査であるということ、また、…（中略）…ことから、標準化回帰係数がどのような傾向にあるのかを見ておくことは一定の意味があると思われる」（お茶の水女子大学, 2018, p. 35）。この「一定の意味」がどの程度であるのか、CTTでは比較できないとされる理論的問題を超えるだけの意味があるか否か、いつそう専門的で緻密な議論が今後必要であろう。

は因果関係はなく、別の因果関係によって、両者が相関することもある。否、保護者の所得の高さと、子どもの学力の高さには、直接的な因果関係はそもそも存在しないはずである。保護者の金銭の獲得が、ただちに子どもの学力の獲得を帰結する、とは考えられない。保護者の所得の高さは、高価な通塾費を負担しても良いという余裕を招き、子どもを通塾させることで、子どもの学力が向上するかもしれない。あるいは、書物を購入したり海外旅行費を負担する余裕を生み、本や異文化に触れることで、子どもの知的好奇心が刺激され、結果として子どもの学力が上がるなど、様々な要因がめぐりめぐって、いわば“磁場”のような複雑な作用圏を形成して、学力を上げていると考えるのが自然であろう（図3参照）。それゆえ各要素の因果的相関のウェイトも含め、この“磁場”の特性を可視化する必要がある。



(田端作成)

図3：SESと学力との相関と因果のイメージ

中村たちによると、観測できない要因には、「時間を通じて変化しない要因」と「時間を通じて変化する要因」とがあり、今回は前者を制御する「固定効果モデル分析」を実施したところ、クロスセクション回帰分析では「世帯所得が学力に対して正の効果を持っているが、固定効果モデルでは世帯所得が学力に与える影響は統計的優位性を失う」結果になった（cf., 赤林, 2016, pp. 78-79）。ここから中村たちは、「これだけをもって『所得が学力に与える影響は無視できる』とまで結論づけるのは早計である」（赤林ほか, 2016, p. 78）としながらも、「クロスセクションデータを利用した分析で示された世帯所得が子どもの学力に与える効果が必ずしも因果関係とはいえず、見かけ上の相関であったことを示唆する結果である」（赤林ほか, 2016, p. 80, 強調は引用者）と結論づけている。

こうした新たな学術的分析結果は、実践への示唆を大いに含んでいる。中村たちはこの結果を受け、生活保護や教育費支援の増額といった「ある時点の世帯所得の格差と学力だけに注目した社会政策・教育政策は、学力格差を縮小するという目的に対しては必ずしも有効性を期待できない」（赤林ほか, 2016, p. 80）とする

13。

13 所得増額が「一時的」ではなく「恒久的」な政策の有効性を、中村たちは示唆している（cf., 赤林ほか, 2016, pp. 80-81, 強調は引用者）。

Ⅲ 「学力」と「非認知能力」

非認知能力が日本の学力研究で本格的に注目されるようになったのは、2015年あたりからである。埼玉県が質問紙調査に非認知能力を加えたのも、2016（平成28）年度だった。世界的には、ヘックマン（Heckman）たちの教育経済学的研究（Heckman & Rubinstein, 2001）が火付け役になった（cf., 遠藤, 2017, p.15）。非認知能力を変数とすることで、諸要因の説明モデルがいつそう複雑多様化することになる。

上記で検討してきたように、学力を「学力テストで測定可能な能力」と割り切った定義をしたとしても、この定義で完結するわけでも、「学力とは何か」の問題に終止符が打たれるわけでもない。むしろこの定義は、実は、学力とは何かに迫る出発点と方向性を定めたのである。学力調査で測定可能な能力は、測定方法が変われば、測定される能力のタイプや質が違ってくる。このことを顕著に示していたのが、CTTとIRTの違いであった。また、測定可能な学力の構成要素に目を向けるならば、この能力は、教育生産関数で表せるような諸要素（変数）の相関関係として考えられる。学力は、諸要素が複雑に影響しあった結果の出力というわけである。なかでも、保護者のSES（社会的経済的地位）が学力に有意に相関することが、従来の多くの研究で実証されてきた。しかも従来の研究では、子どもの学力は保護者のSESでかなりの程度決まるという規定説が優勢であった。ところが、それも近年、観測方法が変わることで、その相関関係の存在そのものが疑問に付されるまでに至っている。さらに無視できない要因として、非認知能力が注目されることで、学力と非認知能力との相関が様々に調査分析され、新たな相関関係が次々と実証されている。この項目が想定されなければ、そもそも観測されなかった（観測上存在しなかった）関係である。

1 「非認知能力」とは何か

非認知能力が加えられることで、観測される関係にどのような変化が現れるのか。これをフォローする前に、そもそも「非認知能力」とは何かを押さえておきたい。

この能力は、英語では「non-cognitive ability（非認知能力）」とか「noncognitive skill（非認知技能）」あるいは「social and emotional competence（社会情緒的コンピテンシー）」ともいわれる。

非認知能力の代表としては、通常、「意欲」「長期的計画を実行する能力」「他人との協調に必要な社会的・感情的制御」などがあげられる（cf., ヘックマン, 2015, p.17）。だが一歩踏み込んで内実を明確化しようとすると、そう簡単ではない。

パネルデータで重要となる「観察（観測）不可能な要因」が、時間的に変動する要因と変動しない要因とに大きく区分されることを考慮すると、非認知能力を時間性の観点から次の3層構造で捉える心理学者遠藤利彦たちの試みも示唆的である。

表3：非認知能力の3層構造

表層	自覚的意識に働きかける訓練・指導等によって、相対的に短期的に、また比較的容易に習得可能と見込まれる構成要素。	（例）個別具体的な社会的スキルやマナー、社会的ルールや規範に関する知識など。
中層	発達の時間軸で徐々に形成され、相対的に安定して持続するが、社会経験や学習の蓄積により変容可能な構成要素。	（例）自尊心、自己効力感、自制心、メタ認知能力、自己内省、他者に対する社会的行動や共感的態度やコミュニケーションスタイルなど。
深層	発達の早期段階にその基盤が形成され、生涯全般を通して相対的に高い一貫性を有し、容易には変化しにくい構成要素。	（例）気質やパーソナリティ、発達早期のアタッチメント（愛着形成）に由来するとされる根源的な自他信頼感など。

深層には遺伝的要素も含まれる。この表からも、非認知能力がきわめて多彩な要素で構成されることがわかる。

非認知能力への注目の火付け役が、ノーベル経済学賞を受賞したヘックマンだったことに象徴されるように、この能力の重要性を顕在化させたのは、経済学だった。経済学ではこれまで、「学力」として測定されるアカデミックな認知能力の高さが、将来の収入や所得、昇進や雇用形態を予測するとされてきた (cf., 遠藤, 2017, p. 7)。これに対しヘックマンたちは、個人の賃金の高低について、高校卒業者、高校中退者、高校卒業と同等の資格保有者を比較したところ、認知能力だけでは賃金の差を説明できず、非認知能力の重要性を示唆した (cf., 遠藤, 2017, p. 7)¹⁴。学力以外の非認知能力が所得など SES の向上につながるというこの示唆は、SES-学力規定説にいつそう大きな風穴を開けることになった。

2 OECD (2015) から

では、学力調査に非認知能力の項目を盛り込むならば、どのような知見が得られるだろうか。

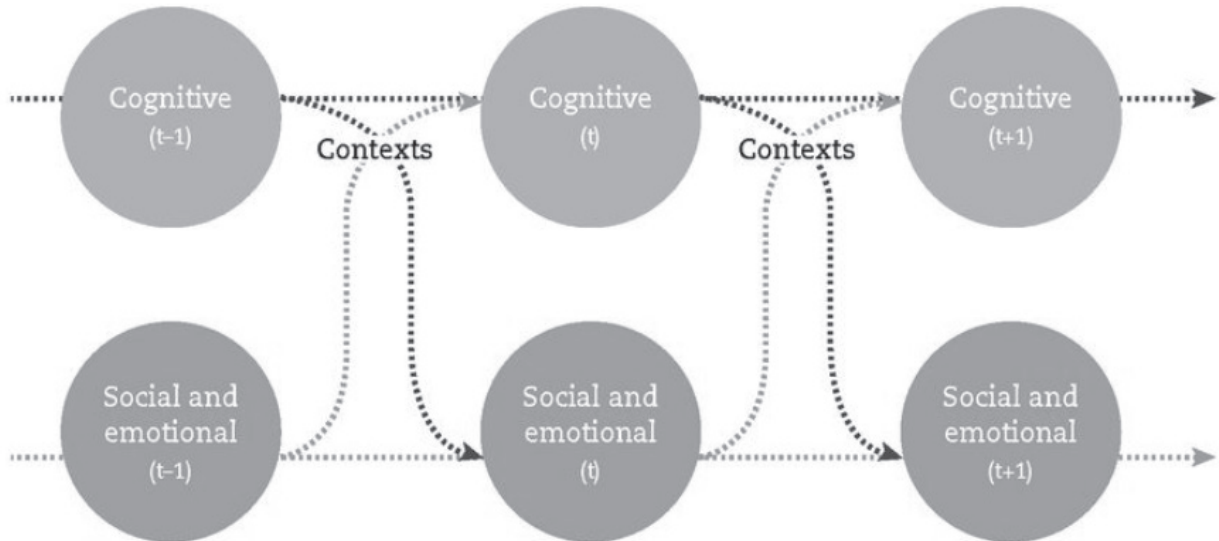
政策決定者や教育関係者に、非認知能力の重要性を実証的に示し、非認知能力を認知能力と同等に家庭・学校・社会で育成する必要性を本格的に提案したのは、OECD の調査報告書「社会を進展させる技能：社会的情緒的技能の力 (Skills for Social Progress: The Power of Social and Emotional Skills)」(2015) である。この調査の核となるリサーチクエスションは、「個人の幸福 (well-being) と社会の発展を牽引する技能とは何か」(OECD, 2015, p. 13) であった。この調査研究で発見された重要な知見のいくつかを、以下箇条書きで紹介しておきたい (OECD, 2015, Executive summary より)。なお以下では、OECD 報告書の「social and emotional skills (社会的情緒的技能)」を「非認知技能」に意識した。

- ・ 非認知技能は、文化的言語的に限定された地域内でなら、信頼できる精度で測定できる。
- ・ 非認知技能は、単独で機能しているわけではなく、認知技能と相互作用し、子どもが将来良い成果 (アウトカム) を達成する可能性を高める。認知と非認知の技能は、経済的社会的アウトカムを向上させるのに有意な役割を演じている。
- ・ 学力などで測定される認知技能の向上は、とくに大学進学を強め、労働市場のアウトカムに強く影響する。一方、我慢強さ、自己評価、社会性などの非認知技能の向上は、健康増進や主観的幸福度の向上と、反社会的行動の抑制に強く影響する。非認知的技能のなかでもとりわけ誠実さ、社会性、情緒的安定は、子どもの将来の展望に強い影響を与える。
- ・ 人生を左右する主要な諸技能は、可変的 (malleable; 影響を受けやすく柔軟) で、とくに非認知技能はそうであり、家庭や教師や社会の介入の意義が大きい。とりわけ幼少期に可変的であるため、幼少期での政策的・教育的介入が効果的である。

¹⁴ ヘックマンが依拠したデータの 하나가「ペリー就学前計画」である。これは 1960 年代にミシガン州で実施されたもので、教育上ハイリスクな子どもとその家庭を対象とした介入プログラムとその効果測定の追跡調査とで構成されていた (cf., 遠藤, 2017, p. 7)。同地区の家庭が実験群と統制群とに分けられ、実験群には幼稚園での教育プログラムが提供され、養育者への養育指導が実施された (cf., 遠藤, 2017, p. 7)。その結果、実験群の子どもは、小学校入学後の数年間、学力で統制群より優秀であったが、小学校中学年では統制群と同等になり、その後両者に学力差は見られなくなった (cf., 遠藤, 2017, p. 7)。ところが 40 年後、実験群だった人々は、犯罪率が低く、年収が高く、持ち家率も高いなど、統制群より良い生活を送っているという結果になった (cf., 遠藤, 2017, p. 7)。ここから、就学前教育の重要性が浮き彫りにされ、それが長期的には学力以外の能力を向上させ、人々のウェルビーイングを高めると仮定されるようになった。

- ・ 諸技能は諸技能を生む (skills beget skills)。

「技能が技能を生む」という知見は、この報告書の一つの結晶点である。図4に表されるように、OECD (2015)によれば、過去の技能は、将来の技能形成の土台となり、その形成には、認知と非認知とのダイナミックな相互作用がある。認知技能と非認知技能は、非常に密接に結びついている (cf., OECD, 2015, p. 39)。



(OECD, 2015, p. 39 より。図表番号は引用者が改め、タイトルは意識した。)

図4：認知技能と非認知技能とのダイナミックな相互形成過程

この相互形成プロセスからすれば、諸技能は「雪だるま式」(OECD, 2015, p. 74)に拡大することになり、「個人が高レベルの技能を持っていればいるほど、後に獲得される技能も大きくなる」(OECD, 2015, p. 39)。

同報告書第4章で詳述される、韓国の長期追跡データをもとにした14歳から15歳への認知・非認知技能のダイナミクス分析は啓発的である。それによると、「14歳時点の社会的情緒的技能の向上…が15歳時点の同技能に与える影響は、14歳時点での同技能のレベルに応じて増加する」(OECD, 2015, p. 74)。つまり、14歳時点の非認知能力が高いほど、15歳時点でのその能力の伸びが大きくなり、結果として非認知能力低レベルと高レベルとの格差は拡大する。そればかりか、「14歳時点の社会的情緒的技能の向上が15歳時点の認知技能…に与える影響もまた、14歳時点の社会的情緒的技能のレベルに応じて増加する」(OECD, 2015, p. 74)。つまり、14歳時点の非認知能力が高いほど、15歳時点の非認知能力だけでなく、認知能力の伸びも大きいという。

また「14歳時点の認知能力の向上が15歳時点の認知能力に与える影響は、14歳時点の認知能力のレベルに応じて増大するが、[非認知の影響力と比べ]より小さい割合である。このことから、将来の認知能力の向上に関しては、現在の認知技能のレベルよりも、現在の社会的情緒的技能のレベルの方が重要である、と示唆される。」(OECD, 2015, p. 74, []内引用者)

もしこれが本当ならば、現在の学力向上よりも、現在の非認知能力向上の方が、将来の学力を向上させるのに効果的だということになる。だが、韓国の14歳から15歳への変化の分析だけでは、こうした一般化には慎重であるべきであろう。しかし、インパクトの大きな分析結果である。

以上からOECD (2015)は、「諸技能の育成は、全体的 (holistic) かつ一貫的 (coherent) でなくてはならない」(OECD, 2015, p. 90)と提言している。

3 埼玉県学力調査から：「非認知能力」「主体的・対話的な学び」「学級経営」との相関

次に、非認知能力を取り入れた埼玉県の調査報告を見てみよう。

2019（令和元）年報告書では、過去3年間のデータをもとに2017（平成29）年度までにおこなった分析結果として、以下の2点をまとめ、それを下記の関係図で表現している。（以下図5まで、括弧内も含め、埼玉県（2019，第2章）からの引用。図表番号のみ引用者が変更。）

- ・ 「主体的・対話的で深い学び」は、児童生徒の「非認知能力」の向上や「学習方略」の改善を通じて、学力を向上させる可能性がある。（①～②）
- ・ 「学級経営」が、「主体的・対話的で深い学び」の実現や、子供たちの「非認知能力」「学習方略」の向上に重要であり、結果として児童生徒の学力向上につながる可能性がある。（⑤～⑦）

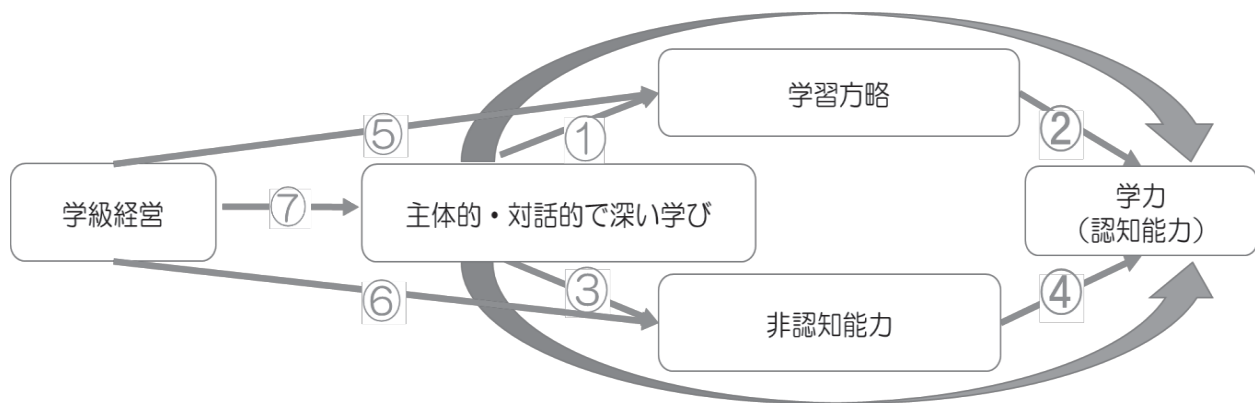


図5：データ分析結果の関係図

この分析は、慶應義塾大学 SFC 研究所に委託されており（埼玉県，2019，第2章）、分析結果の信頼性は高い¹⁵。

この分析では、非認知能力は媒介項となり、主体的・対話的な学びと学級経営を、学力へと関連づける。これまでは知られていなかったダイナミズムである。

対話的な学びと学級経営とは有意な相関がある、との指摘も重要である。対話的学習に積極的に取り組んでいる学級は学級経営が良好であり、またその逆でもある。一見当然とも思える結果だが、これが統計的に実証された意義は大きい。

ただ、代理変数となる質問項目が、対話学習でも学級経営でも、十分ではない点に課題が残る。学級経営と関連する項目は、例えば中学3年生の質問項目では、「学級は落ち着いて学習する雰囲気がありましたか」と「あなたの学級は、いろいろな活動にまとも取り組んでいたと思いますか（体育祭や合唱コンクールなどの学校行事も入ります）」の2つだけである¹⁶。また「主体的・対話的で深い学びの実施」の分析に使用されている項目は、「あなたの〇年生の時の〇〇の授業では、次のようなことがどれくらいありましたか」という形式の質問に対する以下の回答である（埼玉県，2019，第2章，行頭の①～の番号のみ田端が加筆）。

- ① 課題を解決する時に、それまでに習ったことを思い出して解決できたこと
- ② 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

¹⁵ 埼玉県のホームページでは、慶應義塾大学 SFC による詳細な分析報告書「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分析研究」も、平成29年度分と平成30年度分が掲載されている（2020年1月アクセス）。

¹⁶ 「平成31年度『埼玉県学力・学習状況調査』の問題概要」から中学3年生対象の「質問紙調査（PDF）」から、学級経営に関連すると考えられる項目を、田端が抽出した。（ ）内は原文。

- ③ 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
- ④ グループで活動するとき、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
- ⑤ 授業で課題を解決するとき、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
- ⑥ 授業の始めに、先生から、どうやったら課題を解決できるか考えるように言われること
- ⑦ 授業の始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

これら7項目すべてを対話的・主体的な深い学びの指標として合成しているのか、該当しそうな項目だけを抽出しているのかは、報告書では明らかでない。「対話的」な学びに関連しそうなのは、項目④⑤のみである。学級経営でも対話的学びでも、各2項目との相関だけでは、分析結果の妥当性は限定的と言わざるをえない。

同じ埼玉県学力調査をもとにした石川たちの分析では、対話的学習にあたるアクティブラーニングに関して、上記④⑤の項目のみを抽出し、学力との相関を分析している (cf., 石川ほか, 2017, p. 41)。その結果、「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、学力が高い傾向にある」(石川ほか, 2017, p. 40) としている。対話的学びと学力との直接的で有意な相関である。さらに石川たちの分析では、国語もさることながら「算数へのアクティブラーニングの指導の影響が大きい」(石川ほか, 2017, p. 41)。一般に対話的な学習は、正解に到達させる算数・数学には不向きと思われるが、それは思い違いであり、国語よりも数学の学力向上と強い相関が見られるのである。一般的な思い込みに囚われてはならないと教えられる。

対話的学習と学力との有意な相関から、石川たちは次のように指摘する。

「IRT に基づく前年度学力やその他の多くの変数を制御した上でも、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。」(石川ほか, 2017, p. 41) 要するに対話的学びは、国語と数学の学力向上と関連し、とくに数学、とりわけ非通塾層の数学の学力を押し上げる。

以上から対話的な学びは、学力と直接関連するルートと、非認知能力や学習方略を經由して学力と間接的に関連するルートの3経路によって、学力向上と強く関連していると推測できる。この結果からすれば、対話的学習の一つ「子どもの哲学 (philosophy for children : p4c)」は、学力、非認知能力、学級経営、学習方略と有意な正の相関関係があると言える。

4 日本子どもパネル調査 (JCPS) から

IRT とパネルデータによる調査分析は、まだ途についたばかりである。それにしても、これら新たな手法とデータにもとづく近年の分析結果は、認知能力の可変性をいっそう顕在化させ、認知能力の決定要因と形成プロセスの多様性を明らかにし、家庭や学校や社会での教えと学びによる認知能力向上の可能性を示している。それゆえ、地域ごとの家庭や学校や社会での教えと学びによっては、子どもの学力格差は、拡大もすれば是正もされる。日本子どもパネル調査 (JCPS) による、所得と非認知能力との関係についての分析も同じ方向性を示している。

JCPS は、所得-学力の相関に関する先の分析と同様に、所得-非認知能力の相関もまた、クロスセクションデータでの統計的有意性が、パネルデータでは確認できなくなるという驚きの結果を公表している。

非認知能力の指標として JCPS が採用したのは、一つには、保護者対象の「子どもの強さと困難さアンケート」であり、「情緒的不安定さ」「行為問題」「多動・不注意」「仲間関係の持てなさ」の4つの下位尺度の合計点である (cf., 赤林ほか, 2016, pp. 111-112)。もう一つは、子ども対象の「生活の質 (Quality of Life :

以下 QOL) アンケートであり、「身体的健康」「情動的ウェルビーイング」「自尊感情」「家族」「友だち」「学校」の6つの下位尺度の合計点である (cf., 赤林ほか, 2016, p. 112)。分析の結果、「世帯所得が子どもの問題行動・QOLに与える影響は、クロスセクションデータに基づく回帰分析では確認することができたが、パネルデータに基づく固定効果モデルではほぼ確認できなかった。」(赤林ほか, 2016, p. 126) クロスセクション回帰分析では、所得と非認知の正の相関が確認できたばかりか、その相関関係は「年齢とともに強まる」(赤林ほか, 2016, p. 126)。ところが驚くべきことに、固定効果モデルでは、この相関関係が消えるのである (cf., 赤林ほか, 2016, p. 126)。

どうしてこのような矛盾した分析結果になるのだろうか。分析者の赤林英夫たちは、2つの要因を推測している。第一は遺伝的要因であり、第二は時間を通じて変化する脱落変数である。ポイントはいずれも「観測不可能な要因」であり、第一は「時間を通して不変な要因」であり、第二は「時間を通じて変化する要因」である。第二の要因については、JCPSの当該時点ではデータ不足から計測できないという。第一のケースは、クロスセクションとパネルデータの特性を理解する助けになるので、簡単に紹介しておきたい。

赤井たちによれば、「行動遺伝学研究は、認知能力に限らず、非認知能力にも遺伝の影響が顕著にあることを明らかにしている」(赤林ほか, 2016, p. 126)。そうすると、子どもの非認知能力の高さの要因は、親からの「遺伝」と考えられる。一方、親の所得の高さも、遺伝的に親の認知・非認知能力が高いことに起因する。それゆえ、「親の所得」と「子の非認知能力」との間に因果関係がなかったとしても、両者に強く影響する「遺伝的要因」という観測不可能な要因が、両者を共に押し上げているため、両者は見かけ上相関し、観測不可能な固定的要因を制御できないクロスセクションでは、この見かけ上の相関が観測されるが、それを制御するパネルデータではこの相関は消えるわけである。

これは一つの仮説である。しかもこの仮説は、諸能力の可変性を強調する OECD (2015) と整合的ではない。今後の学力論は、遺伝子研究の最新成果にも目配せする必要があるだろう。例えば、この10年の研究では、親のケアが、ホルモンなどの脳からの分泌物に影響するだけでなく、もっと深く、遺伝子発現にも影響を与える証拠が発見されている (cf., タフ, 2019, p. 33)¹⁷。

いずれにせよ、時間的に固定的であろうと可変的であろうと、観測不可能な要因の影響が示唆されればされるほど、それを観測なり観察し、その影響の程度を他の要因と比較検証する必要性が増す。経済学で観察不可能とされる要素を観察し可視化する質的研究 (Qualitative Research) もいっそう重要になるだろう。

おわりに

本稿は、学力をテストで測定される能力と限定することから出発し、測定される学力について、近年注目される議論をレビューした。これを通して、測定・観測される学力が、測定・観測する理論にいかにか依存しているかが明らかになった。これは一種の「観測者問題」(cf., 田端, 2017, pp. 175-176) に他ならない。主要な学力調査のレビューを主とする本稿に、幾ばくかのオリジナリティがあったとすれば、それはこの観測者問題を、学力調査の理論と実践の往還にも見出した点にあるかもしれない。

観測者問題を最初に自覚したのは、現代物理学者だった。観測の仕方を変えることで、光が粒として確認

¹⁷ タフの紹介を引用しておきたい。「マギル大学の研究者らは、母ラットの特定の行動が、子ラットの DNA の配列に起こるメチル化に影響を与えることを明らかにした。子ラットがストレスを受けたときに母ラットが示す温かく繊細な対応、とくにリッキング・アンド・グルーミングと呼ばれるなだめるような行動が、DNA 上で海馬を制御する部位のメチル化を抑制するのだ。海馬は、成長したときにストレスホルモンを処理する部位だ。… (中略) …親の本の小さな配慮が、非常に深いところから——きわめて重要な遺伝子情報に関わる部分まで掘りさげ流用にして——子供の発達を助けるのだ。」(タフ, 2019, pp. 33-34)

されたり、波として確認される矛盾を数学的に整合的に証明したハイゼンベルクの次の言葉が、観測者問題の出発点となった。「観測が事象において決定的な役割を演ずること、そしてリアリティは我々がそれを観測するかしないかによって変わってくる。」(ハイゼンベルク, 2008, p. 30) 観測者問題は、教育研究の領域では、これまで質的研究において自覚されてきたが(田端, 2017, pp. 175-179)、本稿で明らかになったように、それが学力調査でも顕在化してきている。そのきっかけは、調査理論が CTT から IRT へ移行したり、クロスセクションデータからパネルデータに移行するなど、数学的演繹が飛躍的に高度専門化すること(「学力把握の数学化」の飛躍的進展)によってである。

「学力把握の数学化」は、今後なおいっそう進展するだろう。それに合わせて重要になるのは、観測者問題にいっそう自覚的になることである。観測される事象を左右する観測の理論と方法をよく知ることである。確かに、学力を観測する理論は、その数学化の飛躍的進展のため素人の理解をはるかに超えるレベルに達している。しかし、少なくとも本稿がレビューした分析結果は、認知能力と非認知能力とのダイナミズムや、対話的学びによる認知非認知能力の向上など、一定の実践感覚と整合的である。学力測定の素人でも一定の実践感覚のある者たちは、学力とその決定要因に関して、高度な数学に比する把握力を潜在的にもっている(cf., 田端, 2017, p. 176-178)。

そうすると、学力把握の将来の高度な数学化にとって重要になってくるのは、測定の専門家たちと、測定の専門家ではないが実践感覚をもつ人びととのコラボレーションであろう。また本稿では、学力問題に遺伝子研究を参照する必要性も指摘された。それゆえ、多種多様な専門家と非専門家とのハイブリッドな対話(deliberation)の活性化が、今後なおいっそう不可欠である。

引用文献

- 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』有斐閣.
- 遠藤利彦代表, 2017, 「非認知的(社会情緒的)能力の発達と科学的検討手法についての研究に関する報告書」, (accessed Dec. 2019) .
- ヘックマン, J. J., 2015 『幼児教育の経済学』古草秀子訳, 東洋経済新報社.
- ハイゼンベルク, W. K., 『現代物理学の思想』河野伊三郎・富山小太郎訳, みすず書房.
- 石川善樹・伊藤寛武・植村理・田端紳・外山理沙子・中室牧子・分寺杏介・星野崇宏・松岡亮二・山口一大, 2017, 「子どもの能力を計測するための学力テストの現在と展望:エビデンスに基づく教育政策に向けて」RIETI Policy Discussion Paper Series 17-P-010, (accessed Dec. 2019) .
- 川口俊明, 2014, 「国際学力調査からみる日本の学力の変化」『福岡教育大学紀要』63.
- 川口俊朗, 2018, 「PISA から私たちは何を学べるのか? : 『学力調査の設計』という視点から」SYNODOS, (accessed Dec. 2019) .
- 川口俊明, 2019, 「日本の学力研究の動向」『福岡教育大学紀要』68.
- 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.
- 耳塚寛明, 2007, 「小学校学力格差に挑む だれが学力を獲得するのか」『教育社会学研究』80.
- 耳塚寛明編, 2013, 『学力格差に挑む』金子書房.
- 耳塚寛明編, 2014, 『教育格差の社会学』有斐閣アロマ.
- 光永悠彦, 2018, 『テストは何を図るのか: 項目反応理論の考え方』ナカニシヤ出版.
- 中西啓喜, 2017, 『学力格差拡大の社会学的背景』東信堂.
- 文部科学省・国立教育政策研究所, 2017, 「平成28年度 全国学力・学習状況調査 経年変化分析調査」, (accessed Dec. 2019) .

- OECD, 2015, Skills for Social Progress: The Power of Social and Emotional Skills, (accessed Dec. 2019) .
- お茶の水女子大学, 2008, 「青少年期から成人期への移行についての追跡調査 JELS 第 8 集 C エリア基礎年次調査報告」, (accessed Dec. 2019) .
- お茶の水女子大学, 2014, 「平成 25 年度全国学力・学習状況調査(きめ細かい調査)の結果を活用した学力に影響を与える要因分析に関する調査研究」, (accessed Dec. 2019) .
- お茶の水女子大学, 2018, 「保護者に対する調査の結果と学力等との関係の専門的な分析に関する調査研究」, (accessed Dec. 2019) .
- 埼玉県, 2019, 「令和元年度埼玉県学力・学習状況調査報告書」, (accessed Dec. 2019) .
- 田端健人, 2017, 「現象学と臨床教育学: 科学技術への新たな架橋」, 矢野智司・西平直編著『臨床教育学』協同出版.
- タフ, ポール, 2019, 『私たちは子どもに何ができるのか: 非認知能力を育み、格差に挑む』高山真由美訳, 英治出版.
- 東北大学, 2011, 「全国規模の学力調査における重複テスト分冊法適用の試み」, (accessed Dec. 2019) .
- 東北大学, 2018, 「経年変化分析調査との対応づけによる本体調査の年度間比較の試み」, (accessed Dec. 2019) .