

全国学力・学習状況調査の平均正答率をどう受けとめるべきか？

— 「生きられた数値」による〈子ども教育データサイエンス〉の構想—

How to Understand Means of National Scholastic Tests?

— Ideas of 'Data Science of Education for All' based on Lived Numbers —

田 端 健 人
Taketo TABATA

目 次

1. 問題意識と課題設定
2. 都道府県別平均正答率の最上位と最下位との差
3. ヒストグラムと中心極限定理
4. 全国学調スタート時との比較
5. 「学力の全国平均値はほとんど動かない」という仮説
6. アメリカ合衆国の学力平均の長期トレンド
7. 一定の妥当性をもつ3つの仮説
8. 全国学調の受けとめ方に関する提言
9. 教師の実践感覚の精度の高さ
10. 「子ども教育データサイエンス (DS-EFA)」の理論背景としての現象学
11. 数値から実践感覚を問い直す分析結果
12. なぜ「学力」にこだわるのか？

「現象学は世界の神秘と理性の神秘とを開示することを任務としている」
(M. メルロー=ポンティ)

1. 問題意識と課題設定

「うちの市は学力が低いです。全国学調¹平均正答率より○ポイント低く、なかでもうちの学校は全国平均より△ポイントも低いんです。下の下ですよ。」

こうした言葉を、学校や教育委員会をよく耳にしてきた。そのたびに、平均より○ポイントとか△ポイント低いことは、どのくらいの学力差を示しているのか、疑問を感じないではいられなかった。そのポイント差は、由々しき問題なのか、それともさほど気にしなくてもよい問題なのだろうか。平均にして何ポイントの差があるとき、早急に対応すべき問題とみなすべきだろうか。

こうした問題意識から、本稿では、全国学力・学習状況調査の平均正答率をどのように受けとめればよいのか、わたしたちの実感に引き寄せながら検討してみたい²。

¹ 全国学力・学習状況調査の略称。本稿ではこの略称を適宜用いる。

² 当該課題に関する先行研究をCiNiiで検索すると、「全国学力・学習状況調査」では598件だが、「全国学力・学習状況調査+平均」では26件になる（2020年11月12日現在）。この26件の中には、平均正答率の差を話題にしている論文が含まれているが、その差がどれほどのものか、その差の実質的意味を検討している論文はない。

2. 都道府県別平均正答率の最上位と最下位との差

都道府県別の平均正答率で、最上位と最下位はどれほどの差があるのだろうか。平成31（令和元）年度の国語と算数・数学を例に確認しておこう。参考までに全国の平均正答率（公立）も記載しておく。

表1 H31全国学調平均正答率、最上位と最下位³

	小6国語	小6算数	中3国語	中3数学
最上位	秋 田：74.0%	石 川：72.0%	秋 田：78.0%	福 井：66.0%
全国平均（公立）	63.8%	66.6%	72.8%	59.8%
最下位	愛 知：59.0%	北海道：64.0%	沖 縄：68.0%	沖 縄：53.0%
最上位と最下位との差	15.0P	8.0P	10.0P	13.0P

表1からわかるように、最上位と最下位との差は、8～15ポイントである。小6国語の正答率に着目してみよう。100点満点のテストで、74点とった生徒と、59点とった生徒とをイメージするならば、この点差は大きいとも感じられる。「59点しかとれなかった」という印象さえ受ける。ただ、平均が約64点であることを考慮すると、平均と5点差の59点が低いというより、10点差の74点が抜きん出ている印象にもなる。はたしてこうしたイメージや印象で、全国学調の平均正答率を理解してよいのだろうか。

全国学調の分析を専門とする研究者は別として、それ以外の教育関係者にはあまり知られていないかもしれないが、全国学調の採点は、正答数でカウントされている。問題への配点はなく、正答1、誤答0でカウントする。その正答数を全問題数で割り100を掛けた値が正答率である。表1の正答率を正答数で表記すると、表2になる。なお、全国平均（公立）の各欄に当該テストの全問題数を括弧内に併記した。

表2 H31全国学調平均正答数、最上位と最下位⁴

	小6国語	小6算数	中3国語	中3数学
最上位	秋 田：10.4問	石 川：10.0問	秋 田：7.8問	福 井：10.6問
全国平均（公立）	（全14） 8.9問	（全14） 9.3問	（全10） 7.3問	（全16） 9.6問
最下位	愛 知：8.2問	北海道：9.0問	沖 縄：6.8問	沖 縄：8.4問
最上位と最下位との差	2.2問	1.0問	1.0問	2.2問

各学年各教科の全問題数を見て、全10～16間で本当に学力が測れるのか、疑問がわかなくもない。ただ、問題数が多くなると、児童生徒の解答に時間とエネルギーが費やされ、集中力がもたなくなり、かえって学力が測定できなくなる恐れもある。学力測定の専門家が作成したテストなのだから問題数は妥当であろう、と仮定しておく。

最上位と最下位との正答数の差は1～2.2問である。例えば中3国語なら、最下位の自治体の生徒全員が1問多く解答できれば、最下位から最上位に躍進する。小6国語でも、最下位の自治体の子どもが平均で0.7問多く正答すれば、全国平均に並び、2問多く正答できれば、トップクラスに躍り出る。このように見ると、

³ Webサイト「Education Career」で都道府県別ランキングを確認し、最上位と最下位の自治体について国立教育政策研究所Webサイト「平成31年度（令和元年度）全国学力・学習状況調査【都道府県別】および【指定都市別】調査結果資料」より当該自治体の平均正答率を記載した。

⁴ 図2の自治体について、注3で利用した「平成31年度（令和元年度）全国学力・学習状況調査【都道府県別】および【指定都市別】調査結果資料」より、「平均正答数」を転記した。

最上位と最下位との差はさほどではない気がする。全国学調の結果が発表されるたびに、メディア等で都道府県別に1位から47位までランクづけされる。そのため、1位とか上位は喜び、下位はショックを受ける。しかし、1位と47位の差は、1～2.2問の正答数の差なのである。

都道府県別の平均で、最上位と最下位との差は大きいのか、小さいのか。どちらの印象が正しいのだろう。統計的な手法で吟味してみよう。

3. ヒストグラムと中心極限定理

平成31年度で最上位と最下位の平均正答数の差が最も大きかった小6国語を、改めて取り上げよう。最上位と最下位ならびに全国平均の正答数分布グラフ（ヒストグラム）は図1の通りである。

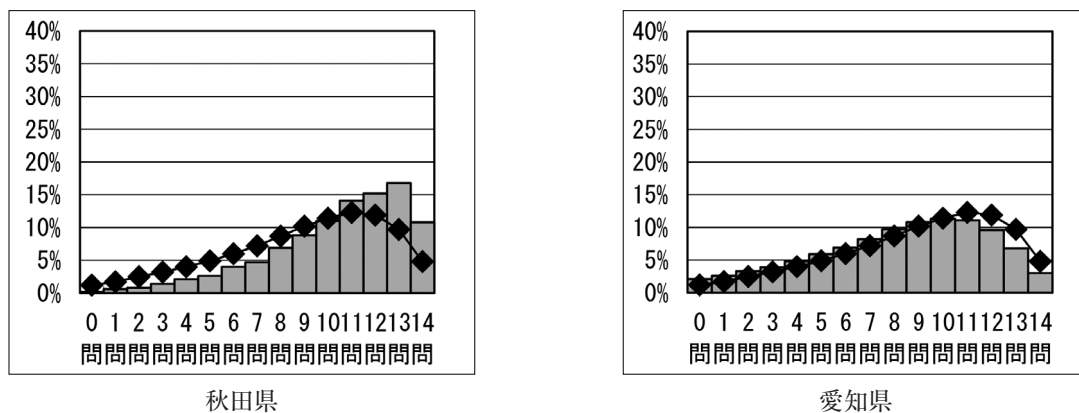


図1 H31全国学調小6国語、最上位と最下位の正答数分布グラフ⁵

横軸は正答数、縦軸は児童の割合、◆は全国平均である。この図から秋田県は9問以下の正答者の割合が全国平均より少なく、11問以上の正答者の割合が全国平均より多いことがわかる。対して愛知県は、9問以下の正答者の割合が全国平均より若干多く、11問以上の正答者の割合が全国平均よりも少ない。しかし、横軸の目盛が5%刻みと小さいことを考慮すると、全国平均との差はわずかに見える。

加えて、都道府県別の平均値は似たり寄ったりになる、という見解を支持する統計上の現象がある。「中心極限定理」である。「中心極限定理」とは、「標本を抽出する母集団が平均 μ 、分散 σ^2 の正規分布に従う場合においても、従わない場合においても、抽出するサンプルサイズ n が大きくなるにつれて標本平均の分布は『平均 μ 、分散 σ^2/n 』の正規分布に近づく」(cf., 石村, p.90 : 小島, p.79, p.141) という定理である。各都道府県を「標本」とみなせば、それらの平均は母集団である全国の平均に近づくと考えられる。ただ「サンプルサイズ n 」がどのくらいの人数になれば、中心極限定理から予測できる結果になるか、疑問は残る。

当該自治体や全国の集計児童数を見てみよう。小6国語の秋田県の集計児童数は6,960名、愛知県は67,141名で、およそ10倍の差がある。中心極限定理にもとづくなら、愛知県の平均の変動（標本誤差）は秋田県のそれと比較して約3分の1程度小さくなる、と見積もることができる。ちなみに全国の集計児童数は、1,028,203名である。

これらのヒストグラムは、どのくらい近似しているのだろうか。2つの標本について、それぞれの母集団

⁵ 国立教育政策研究所Webサイト「平成31年度（令和元年度）全国学力・学習状況調査【都道府県別】および【指定都市別】調査結果資料」より転載。

の確率分布が一致しているかどうか、コルモゴロフ=スミルノフ検定 (K-S検定) にかけてみよう。フリーウェアRで検定したところ、全国分布と秋田分布と愛知県の同等性について、表3の結果となった⁶。

表3 小6国語の全国分布と秋田分布と愛知分布のK-S検定結果

全国分布と秋田分布の同等性	D = 0.2, p-value = 0.925
全国分布と愛知分布の同等性	D = 0.133, p-value = 0.999
秋田分布と愛知分布の同等性	D = 0.2, p-value = 0.938

帰無仮説 (H_0) は、「全国分布と秋田分布は等しい」である。有意水準5パーセントでこの帰無仮説を棄却するためには、p値 (p-value) が0.05未満 ($p < 0.05$) となる必要がある。K-S検定の結果、 $p = 0.925$ なので、 H_0 は棄却されない。それゆえ統計学的には、対立仮説 (H_1) 「全国分布と秋田分布は等しくない (=両者には差がある)」が正しいとはいえない。統計学的に言えるのはここまでだが、後述する「子ども教育データサイエンス」からすれば、全国分布と秋田分布は同等とみなしてよい。

同様に「全国分布と愛知分布は等しい」、また「秋田分布と愛知分布は等しい」とする帰無仮説のK-S検定でも、表3の2段目と3段目の通り、p値はいずれも $p < 0.05$ を満たさず、両仮説は棄却されない。結果、全国分布、秋田分布、愛知分布は、すべて同等とみなしてよい。

煩雑になるので表記はしないが、小6算数、中3国語、中3数学について、同様にK-S検定したところ、 $p < 0.05$ で帰無仮説が棄却されるケースはなかった⁷。

4. 全国学調スタート時との比較

では、中心極限定理から予測される平均の変動幅を考えると、全国平均と最上位と最下位との差は、どの年度でも実はそれほどではなかったのだろうか。全国学調がスタートした平成19 (2007) 年度を見てみよう。差が出やすいB問題に注目しよう。

表4 H19全国学調正答数 (平均正答率)、最上位と最下位⁸

	小6国語B (全10問)	小6算数B (全14問)	中3国語B (全10問)	中3数学B (全17問)
最上位	秋田：6.9問 (69.0%)	秋田：9.6問 (68.6%)	秋田：7.7問 (77.0%)	福井：11.5問 (67.6%)
全国平均 (公立)	6.2問 (62.0%)	8.9問 (63.6%)	7.2問 (72.0%)	10.3問 (60.6%)
最下位	沖縄：5.3問 (53.0%)	沖縄：7.6問 (54.3%)	沖縄：6.4問 (64.0%)	沖縄：8.1問 (47.6%)
最上位と最下位との差	1.6問 (16.0P)	2.0問 (14.3P)	1.3問 (13.0P)	3.4問 (20.0P)

⁶ Rスクリプトについては、本稿末尾の「資料」を参照いただきたい。

⁷ 「資料」のRスクリプトを参照いただきたい。

⁸ 国立教育政策研究所Webサイト「平成19年度 全国学力・学習状況調査【都道府県別】集計結果」をもとに田端が作成。

最上位と最下位との差は1.3～3.4問、正答率で13～20Pである。平成31年度と比較すれば、最上位と最下位との差は、正答数（カッコ内は正答率）において、平成19年度で1.3～3.4問（13～20P）、平成31年度で1～2.2問（8～15P）である。その差に違いを見出すのが難しい数値である。注意深く見れば、最上位と最下位の平均の差は、正答数では平成19年度は3.4（中3 数学B）マイナス1.3（中3 国語B）で2.1問、平成31年度は2.2（中3 数学）マイナス1（小6 算数と中3 数学）で1.2問であり、全国学調スタート時より平成31年度は、都道府県の平均値の差がいつそう縮まったようにも見える。ところが、正答率にすれば、最上位と最下位との差は、平成19年度も平成31年度も7Pと同じになる。やはり全国学調スタート時から、最上位と最下位の平均の差はそれほど大きくはなかった。中心極限定理から予測できる都道府県間の平均の変動幅の小ささは、スタート時からすでに予見できた。

念のため、差が最も大きい中3 数学Bを、K-S検定にかけてみた。結果は表5の通りである⁹。

表5 H19 中3 数学Bの全国分布と福井分布と沖縄分布のK-S検定結果

全国分布と福井分布の同等性	D = 0.167, p-value = 0.963
全国分布と沖縄分布の同等性	D = 0.111, p-value = 0.999
福井分布と沖縄分布の同等性	D = 0.222, p-value = 0.766

帰無仮説（ H_0 ）は、「それぞれの分布は等しい」であり、有意水準5パーセントでこの帰無仮説を棄却できるケースはなかった。やはり全国学調スタート時から、各都道府県の正答数分布相対度数に、有意な差はなかった。

誤解を招かないために、ここで次のことを強調しておきたい。

全国学調がスタートし、それへの賛否が渦巻くなか、各自治体や各学校は大なり小なり学力向上に取り組んできた。そうした努力と成果を、本稿は決して否定するものではない。最上位と最下位の差は、スタート時も12年後も実質的にはないとの指摘は、各自治体の学力向上の成果を批判しているわけではなく、学力向上の取り組みの成果は平均値では見えにくいことを意味している。

5. 「学力の全国平均値はほとんど動かない」という仮説

サンプルサイズが大きいとヒストグラムは正規分布に近づく、という中心極限定理の通りならば、「全国の児童生徒の学力の平均は、経年でほとんど変動しない」という大胆な仮説も成り立つことになる。

しかも、この仮説を支持する複数のエビデンスがある。

一つは、PISAの調査結果である。PISAは項目反応理論（IRT）で設計されており、全国学調とは異なり、調査対象集団が異なっても、年度をまたいで、学力の経年比較が可能である（cf. 数実, p.204）。

国立教育政策研究所2020「OECD 生徒の学習到達度調査2018年調査（PISA2018）のポイント」¹⁰によれば、日本の「数学的リテラシー及び科学的リテラシーは、…調査開始以降の長期トレンドとしても、安定的に世界トップレベルを維持している」とOECDは分析した。これは日本の平均が世界のトップレベルを維持し、さほど動かないことを意味する。また「読解力」は、「前回より平均得点・順位が統計的に有意に低下」としながらも、「長期トレンドとしては、統計的に有意な変化が見られない『平坦』タイプ」とOECDは分析している。「平均が前回より有意に低下」したのは、「その分野が中心分野でなかった年の分野内の対象範囲

⁹ このRスクリプトは、「資料」のTable 2に記載した。

¹⁰ https://www.nier.go.jp/kokusai/pisa/pdf/2018/01_point.pdf（最終閲覧日2020.11.12）

の縮小に関連している可能性を示唆している」とされ、「それぞれの分野が中心分野であった年（読解力については2000年、2009年、2018年、科学的リテラシーについては2006年と2015年）だけを考慮すると、結果はより安定していた」とされる¹¹。つまり読解力でも、日本の平均は経年で安定しているのだ。それゆえ、平均スコアの2000年～2018年の長期トレンドに関するOECDの分析によれば、日本は、「平均得点のトレンドに統計的に有意な変化がない」「平坦」タイプである。

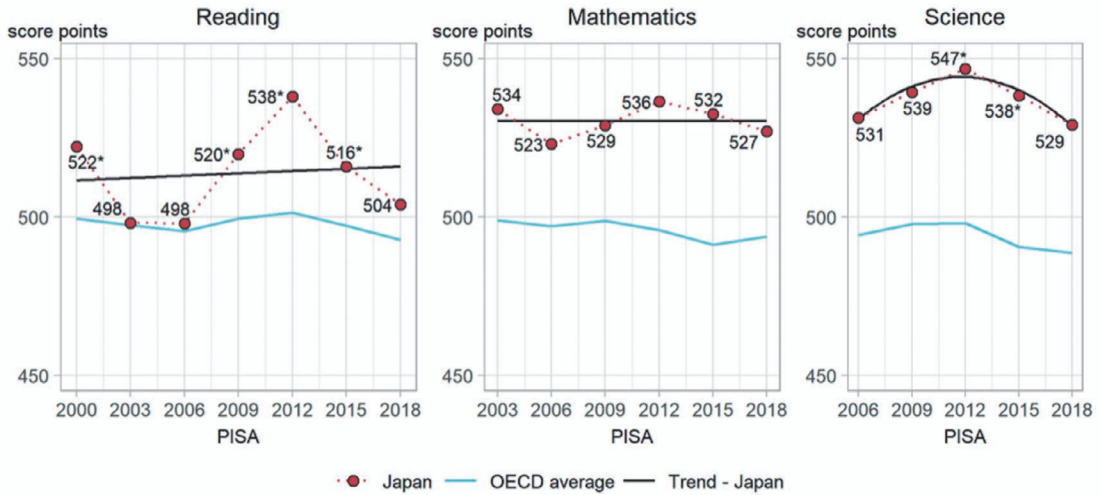


図2 読解力、数学的リテラシー、科学的リテラシーの傾向¹²

ただPISA2018では、「平坦」タイプの他に、「上昇」や「下降」など変動するタイプがある。このことからすれば、たとえ中心極限定理が働くにしても、学力スコアの国単位の平均値は変動することもある。見方を変えれば、IRTで設計されているPISAは、わずかな変化を測定できるほど、精巧に作られているともいえる。ただし、国単位の平均値の変動は、18年という長期スパンでないと見えない変動である。このことは、裏を返せば、国の平均の変動幅は1年とか5年単位では見えないほどわずかである、ということでもある。

長期トレンドに変動タイプがあるとしても、平坦タイプではとりわけ、一国の児童生徒の学力の平均スコアは動きにくい。平坦タイプに分類されたのは、日本の他、フランス、カナダ、アメリカ合衆国、イタリア、イギリスである¹³。これらの一つ、アメリカ合衆国（以下、「米国」）の調査では、学力の平均値が動かないことを示す別のエビデンスもある。

6. アメリカ合衆国の学力平均の長期トレンド

連邦教育達成度調査（National Assessment of Educational Progress: NAEP）による読解力と数学的リテ

¹¹ OECD「Country Note Programme for International Student Assessment (PISA) Results from PISA 2018」pp.3-4. (https://www.oecd.org/pisa/publications/PISA2018_CN_JPN_Japanese.pdf（最終閲覧日2020.11.12））

¹² OECD「Country Note」p.3より転載。

¹³ 国立教育政策所2020「OECD 生徒の額種到達度調査2018年調査（PISA2018）のポイント」「2. 読解力について」を参照。

ラシーの平均スコアを見てみよう。それによると、平均スコアは、1971年から2012年までの期間で、スコア0～500のレンジで、図3、図4のようにわずかな変動しか示さない。

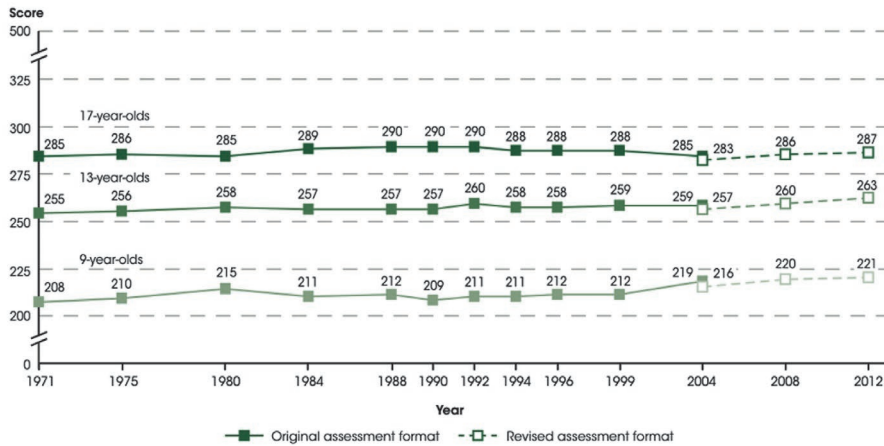


図3 読解力平均、1971年から2012年までの長期トレンド (NAEP)¹⁴

例えば変動幅が大きそうに見える9歳児でも、最低208ポイント、最高221ポイントでその差はわずか13ポイント、全体のレンジに対して2.6%しか変動していない。



図4 数学的リテラシー平均、1971年から2012年までの長期トレンド (NAEP)¹⁵

数学的リテラシーの9歳児平均でも、最低219ポイント、最高244ポイント、その差25ポイント、スコア0～500のレンジの5.0%の変動幅しかない。

ところが、米国でも、子どもや若者の学力低下が、国家存亡の危機として議論された時期があった。そのハイライトは、1983年4月、当時の連邦教育省長官ベルに答申された報告書「危機に立つ国家 (A Nation

¹⁴ https://nces.ed.gov/programs/coe/pdf/coe_cnj.pdf (最終閲覧日2020.11.12)

¹⁵ https://nces.ed.gov/programs/coe/pdf/coe_cnj.pdf (最終閲覧日2020.11.12)

at Risk)」である。「我々の国家は危機に瀕している。我が国がかつて絶対的優位を誇った商業、工業、科学、技術革新は、世界中の競争相手によって取って代われようとしている。」¹⁶このように始まる同報告書は、米国の教育水準低下に警鐘を鳴らす。教育水準低下を示す指標が13項目あげられ、それらのいくつかで学力テストの平均値が話題にあげられている。その一つ、「大学進学適性試験 (Scholastic Aptitude Tests: SAT)」への言及に注目しよう。同報告書は、次のように危機感をあらわにしている。

「カレッジ・ボードの大学進学適性試験 (SAT) は、1963年から1980年にかけて実質的に途切れのない低下を示している。平均言語スコアは50ポイント以上の低下、平均数学スコアは40ポイント近く落ちている。」¹⁷

National Center for Education Statistics (NCES) の統計をもとに、1967年¹⁸から2016年までのSATの平均言語スコアと平均数学スコアをグラフ化してみよう¹⁹。

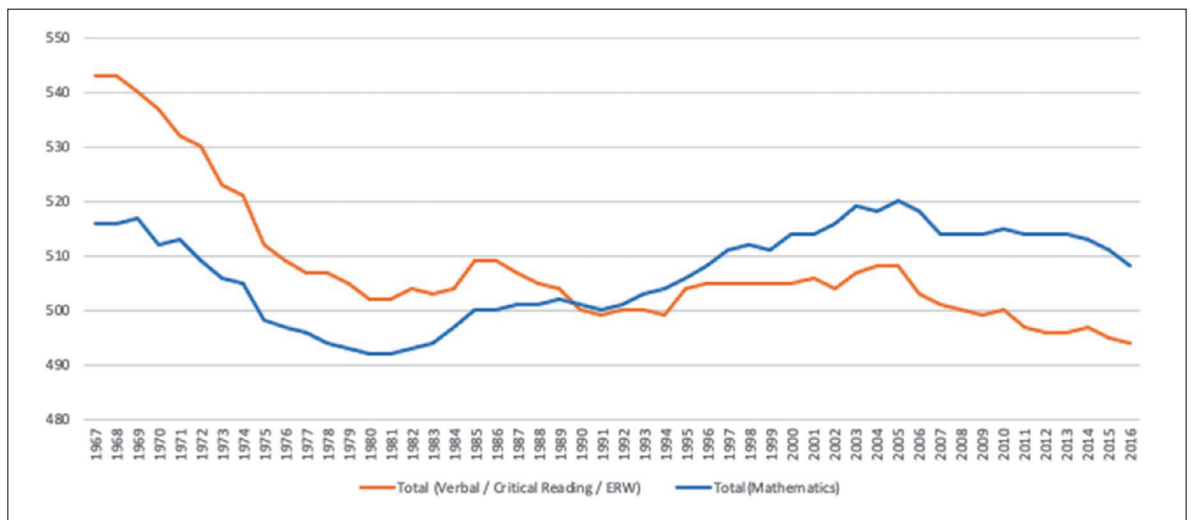


図5 SATの平均言語スコアと平均数学スコア (1967-2016)

確かに報告書が指摘するように、1967年から見ても1980年にかけて、言語スコアも数学スコアも急落して

¹⁶ The National Commission on Excellence in Education, 1983, "A Nation at Risk: A Report to the Nation and the Secretary of Education United States Department of Education", A Nation At Risk. PDF版にはページ番号が付されていないため、節のタイトルを記す。

https://edreform.com/wp-content/uploads/2013/02/A_Nation_At_Risk_1983.pdf (最終閲覧日2020.11.20)

¹⁷ 前掲, Indicators of the Risk.

¹⁸ 1963年まで遡りたかったが、見つけることができたデータは、1967年からだった。

¹⁹ Web上の“Average SAT Scores of College-Bound Seniors (1952-present)” (PDF) の数値をもとに田端がエクセルで作図した。

<https://www.erikthered.com/tutor/historical-average-SAT-scores.pdf> (最終閲覧日2020.11.20)

なお、1967年から1986年までのデータは、その後の新しいスケールに換算されている。そのため、作図で用いたスケールでは、「危機に立つ国家」で指摘される低下ポイントと同じではない。

いるように見える。しかし、図3、4のNAEPのデータでは、1971年から1980年の17歳の学力スコアでは、このような急降下は見られなかった。こうしたデータの齟齬は何に起因するのだろうか。また、数学スコアは、1980年ごろを境に上昇傾向にある。これは同報告書をきっかけとした教育改革の効果のようにも見える。

しかし、図5にはグラフのマジックがある。2016年3月以前のSATは、各教科の最低が200点、最高が800点である。縦軸の最低を200、最高を800として同じデータを作図し直すと図6になる。

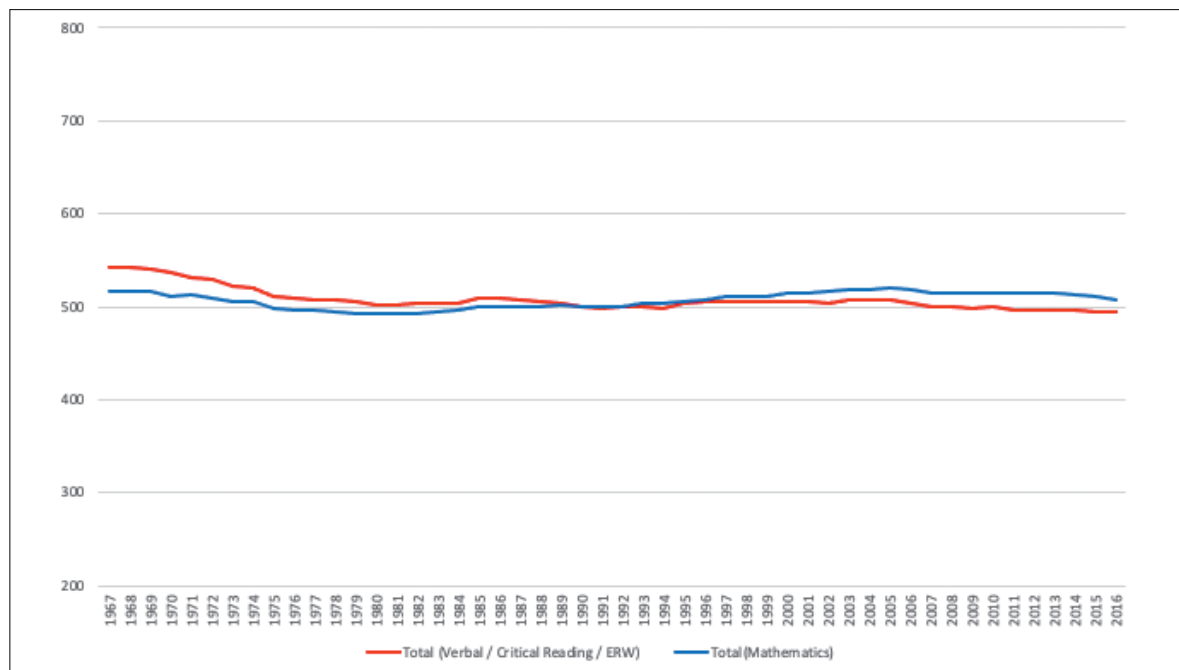


図6 SATの平均言語スコアと平均数学スコア（1967－2016）（縦軸目盛を変更）

NAEPのデータとの齟齬は、こうしたグラフのマジックにあったようだ。同報告書が「危機」の一指標とするSAT平均の低下は、図6では、およそ10年をかけて生じた非常に緩やかでわずかな低下であった。図6で見れば、国レベルの平均はほとんど動かないという仮説が説得力をもつ。

仮に、1963年から1980年にかけて米国が危機に立っていたとすれば、図6のグラフ、つまりSATの平均を中立的に表現するグラフは、その危機を表現するには適切でないことがわかる。国家レベルの平均は、大きな社会現象や国家政策を表現するには、適切ではない。

ちなみに、同報告書をきっかけとして、米国では様々な教育改革が実施されたが、図6の通り、SATの平均言語スコアが1967年の水準に回復した形跡はない。1980年以降、米国の教育は危機に立ち続けているのだろうか。

7. 一定の妥当性をもつ3つの仮説

以上の全国学調やNAEPやSATの学力テストの平均値の長期トレンドから、次の仮説には一定の妥当性がある。

【仮説1】 米国や日本の児童生徒の学力の平均は、この20年間ほとんど変化していない。

これは、米国や日本の児童生徒というサンプルサイズと中心極限定理、ならびにPISAやNAEPやSATなどで得られた経験的なエビデンスを踏まえての仮説である。

さらに本稿2節の都道府県別の最上位・全国平均・最下位の分布の同等性の検定結果等を踏まえて、次の仮説を置くことができる。

【仮説2】 都道府県間の分布に有意な差はない。

さらに中心極限定理と経験的なエビデンスを踏まえると、以下の仮説を引き出すことができる。

【仮説3】 サンプルサイズが7,000名程度を超えると、平均差に実質的な意味はなくなる。

7,000名という人数は、本稿2節で言及した秋田県の集計児童数6,960名をもとにしている。正確には、集計児童がさらに少人数の県もあるので、仮説2から導かれる仮説3のサンプルサイズはいっそう小さくてもよく、5,000名程度としてもよいかもしれない。

8. 全国学調の受けとめ方に関する提言

中心極限定理や経験的なエビデンスや都道府県間の分布の検定から引き出される以上の仮説から、全国学調の受けとめ方に関して、以下のことを提言したい。

第一に、各都道府県の平均正答率を比較しその高低を論じる言説を控えること。「X県は全国で○位」とか「Y県は全国平均を超えた」とか「Z県の学力を全国平均レベルにどう高めるか」といった言説は、もうやめた方がよい。都道府県ごとの平均差に実質的な意味はなく、どの都道府県の分布も全国分布と同等なのだから、その差をことさらあげつらうのは、「序列化」や「過剰な競争」を煽ることにしかならない。

その代わりに、第二に、都道府県の平均の最高点と最低点を、当該学年の当該教科の学力の「平均値帯」(生硬な造語だが)と捉えることを提案したい。例えば、平成31年度全国学調、小6国語でいえば、平均正答率59~74%を、全国の小学6年生の平均値帯とみなすことである。そして第三に、現場の先生方は、例えば平成31年度小6国語なら、平均正答率を59~74%の幅でとらえ、全国学調の各児童の個票を、日々接する子どもの学力を把握する目安として活用いただきたい。例えば、6年生のA君の国語の正答率が59%程度であったなら、「A君は全国の平均的な学力を身につけている」とみなしてよい。あるいは、B小学校の6年1組の国語の平均正答率が74%であったなら、「全国トップクラス」とみなすのではなく、「全国の平均的な学力をもつ集団」とみなすのが妥当である。

さらに第四に、現場の先生方には、目の前の個々の子どもの日頃の学習状況や達成度についての、教師のいわゆる直感的経験的理解(実践知=フロネーシス)をもとに、全国学調の個票の正答率を見積もることを提案したい。例えば、小6のC君の国語の学習状況や達成度について、担任のX先生や教務主任のY先生が「かなり課題がある」と感じていたとする。そして、全国学調の国語のC君の正答率が40%であったとしたなら、全国学調の正答率40%という数値を、「かなり課題がある」ことを示す数値と捉える、という手法である。もしもそのC君が全国学調国語で59%前後の正答率であったなら、C君の頑張りや成長を高く評価してよい。つまり、目の前の子どもの実態(学力)についての教師の実践感覚によって、全国学調の正答率の数値を意味づけるやり方である。こうすることによって、全国学調の数値が、生きた数値に受肉される。これが「生きられた数値」によるデータサイエンス、名づけて「すべての子どものための教育データサイエンス(Data Science of Education for All)」、略称「DS-EFA(ディーエス・イーファ)」あるいは「子ども教育データサイエンス」への第一歩である。

9. 教師の実践感覚の精度の高さ

多くの教師が感覚的に捉えていること（実践感覚や実践知）は、ときに高い精度をもつ。このことは、児童生徒の学力についての教師たちの実践知にも当てはまる。それを示す好例を、ここで紹介しておきたい。柴山直教授たちの「東日本大震災の学力への影響」（東北大学、2014）の調査研究である。2011（平成23）年3月の東日本大震災で津波被害を受けた地域の児童生徒は、肉親を失ったり、家や財産を失ったり、仮設住宅住まいだったり、長時間のスクールバス通学だったり、学校も間借りだったりした。この調査研究は、深刻な被災が児童生徒の学力にどのような影響を与えたかを、IRT（項目反応理論）の推算値による経年比較によって明らかにする、という高度で精緻な分析である。

本稿でフォーカスしたいのは、C市のC中学校とC2中学校のIRT推算値と、現地での聞き取り内容との驚異的な一致である。まずIRT推算値を転記する（図7）。

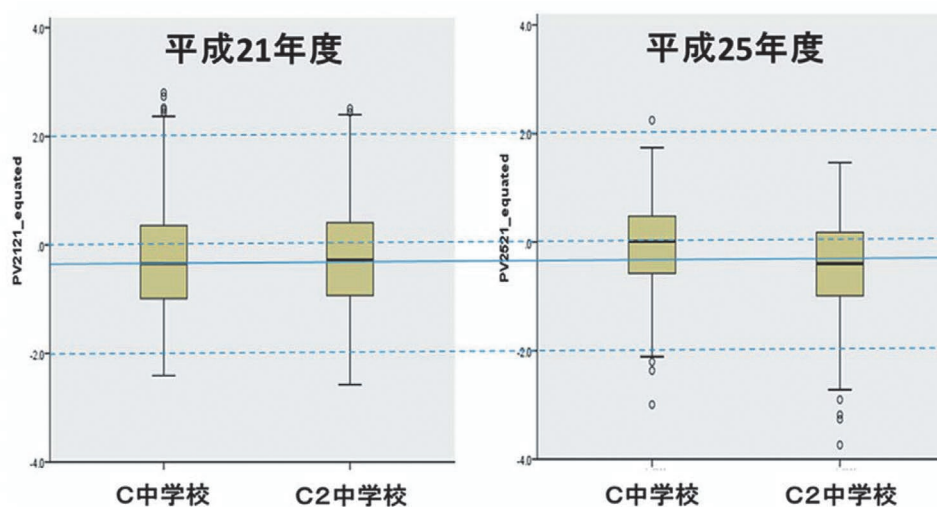


図7 C中学校とC2中学校のIRT推算値の箱ヒゲ図²⁰

柴山教授の情報提供によると、C市教育委員会でのヒアリングで、「C中学校とC2中学校とは震災前は同じ学校規模で同じ学力だったが、震災後、C中学校は学力が上がり、C2中学校は学力が下がった」との情報を得た。IRT推算値で比較したところ、学校が違えば箱ヒゲ図がズレるのが普通であるにもかかわらず、震災前の平成21年度の両校の推算値は、図7左のようにほぼ一致していた。この事実には柴山教授は驚いたという。そして、C市教育委員会など、現場の実践感覚の精度の高さを実感したともいう。視点を変えるなら、こうした実践感覚を数値化できるほど、IRT推算値の精度が高いともいえる。全国学調は古典的テスト理論（CTT）による設計であり、経年比較ができないが、IRTは経年比較ができるいっそう高度な統計学であり、テストの精度が上がることによって、専門家である教師の実践感覚にいっそう肉迫するようになっていく。

図7を、柴山教授たちによる聞き取り調査と合わせて読みとってみよう。こうした読みとりも「子ども教育データサイエンス（DS-EFA）」の手法である。図7からは、C中学校が震災2年後の平成25年度に、被災以前よりも学力を上げていることが一目瞭然である。学力どころではないと思われるあの状況下で、学力が向上したのは、驚きの他にない。対して、C2中学校は、学力を下げている。学力の最大値が大きく下がり、

²⁰ 東北大学, p.48.

最小値も見てわかるほど下がっている。しかし、中央値の低下は見た目上はごくわずかである。教育委員会でのヒアリングと合わせるなら、教育委員会の関係者は、この変化を感覚的に非常に鋭く捉えていたことになる。また、C中学校長からの聞き取りによれば、C2中学校は500名から1,000名規模の避難所になり、「学校にはふさわしくない、大人達のトラブルを生徒達が直接見てしまうことが多々あった」(東北大学, p.56)という。「その結果、避難所が解消された後、一部の生徒達が10月以降暴走しはじめ、教師がそれを止めようとしても『なぜ大人達がやっていたのに自分達がルールをやぶってはいけないのか』として、静止がきかなくなった。校舎の中を荒らすようなことはなかったが、そのため、授業がきちんと成立しない時期が確かに存在した」(前掲書, p.56)という。これほどの実情から改めて数値を読み直すなら、中央値の低下がごくわずかでしかないことに、かえって驚いてもよい。「暴走」した10月が何年度かは記されていないが、この時の荒れを教師と生徒たちがなんとか立て直そうと踏ん張ったことが、学力の中央値を下支えしたと読みとれる。

10. 「子ども教育データサイエンス (DS-EFA)」の理論背景としての現象学

前節で紹介した調査研究(東北大学, 2014)は、学力の経年変化から聞き取り対象の候補を絞り込み、聞き取りと推算値を重ね合わせるという、数量的研究と質的研究との効果的な組み合わせである。DS-EFAは、こうした調査研究方法に学び、数量的データを教師の実践感覚で受肉する試みである。DS-EFAは、教師の実践感覚を数値で裏づける試みであり、逆にまた、数値によって教師の実践感覚を裏切り、教師の児童生徒理解をいっそう深める試みでもある。

この構想の理論背景には、現象学がある。統計学もまた科学の一つであるが、現象学は、科学技術的世界と生活世界とを区別し、その関係を解明してきた。メルロー＝ポンティの次の言葉も、この意味で現象学的である。

「事物そのものへとたち帰るとは、認識がいつもそれについて語っているあの認識以前の世界へとたち帰ることであって、一切の科学的規定は、この世界にたいしては抽象的・記号的・従属的でしかなく、それはあたかも、森とか草原とか川とかがどういふものであるかをわれわれにはじめて教えてくれた〔具体的な〕風景にたいして、地理学がそうであるのとおなじことである。」(メルロー＝ポンティ, p.4, [] 内邦訳書)

「認識以前」の知覚的世界つまり生活世界と、「科学的規定」つまり科学技術的世界との差異を理解し、両者を明確に区別することを、メルロー＝ポンティから学ぼう。わたしたちが実際に目にし、音を聞き、匂いを嗅ぎ、そよ風を感じる現実風景と、地図などのような、その風景を地理学的に処理した科学的世界との違いを理解しよう。

ただ、科学技術的世界が生活世界に対して「抽象的・記号的・従属的」にとどまるか、については議論の余地がある。この二つの世界の関係は、今後いっそう詳しく究明される必要がある。例えば、生活世界の方が、科学技術的世界に対して「従属的」となることもある。すでにフッサールは、この事態をいち早く見抜いていた節(ふし)がある。

「すでにガリレイのもとで数学的な基底を与えられた理念性の世界が、…わたしたちの日常的な生活世界の下に、こっそり差し込まれていること(Unterschiebung; 押しつけ)は、きわめて重要なこととして注意されねばならない。この密かな差し込みは、直ちにその後継者たちによって、…相続されることになった。」(Husserl, S.48f. =フッサール, p.69)

『数学と数学的自然科学』という理念の衣は、…一切を包括し、〔理念の衣をまとった〕これら一切は、科学者たちと同じく教養人たちにとって、『客観的に現実で真実の』自然として、生活世界を代理し、それを変装させる (verkleiden)。』(Husserl, S.52 =フッサール, p.73, [] 内引用者)

数学的自然科学の世界が、生活世界の土台に挿入され、生活世界を代理したり変装させる、との洞察である。そうならば、生活世界の方こそ「従属的」になる。ハーバーマス流に言えば、システムによる生活世界の隷属化・植民地化である。例えば、行ったことのない街の有名レストランをめざして車を走らせるとき、わたしたちはGoogle Mapやカーナビゲーションを頼りにする。いずれも人工衛星からの位置情報をもとにした科学技術的世界である。わたしたちにとっては、カーナビやiPhoneの画像が主であり、眼前に展開するはじめて見る風景は従である。わたしたちは普通、画像のルート案内の方を眼前の風景よりも信頼する。「えっ、こんな細い道に入るの？ 本当？」と疑うこともあるが、たいがいは、わたしたちの実感より、ルート案内の方を「本当」とみなし車を走らせ、結果的に目的地に到着し、ルート案内が「本当」だったことを実感する。この場合、有名レストランのあるこの街を、わたしたちに「はじめて教えてくれる」のは、現実の風景ではなく、科学技術である。

こと「学力」については、両者の関係はいっそう複雑に絡み合っている。ここでは「学力とは何か」の議論には立ち入らない。多くの場合、教育委員会や学校長が「学力向上」で問題にするのは、全国学調や民間学力テストで測定され、平均値や偏差値へと統計学的に処理された数値である。それゆえ、この「抽象的・記号的」な数値が独り歩きすることも少なくない。DS-EFAは、こうした数値の独り歩きにストップをかけることに挑む。科学技術的世界と生活世界とを区別する現象学は、こうした数値の独り歩きに対して、わたしたちの感度を高めてくれる。

さらに、DS-EFAは、その数値が引き出されたおおもとの個々の子どもの「学力」へと「たち帰る」。この意味でも、それは現象学的である。個々の子どもの「学力」を最もよく知る人物の一人は、その子どもたちと日々接している現場の教師たちである。個々の子どもの学力についての現場教師の実践知はというと、授業中の子どもの言動や取り組み、宿題への取り組み、テスト結果など、多種多様なデータによって構成されている。多種多様な感覚や経験値や数値で織り上げられた教師の実践知から、学力の数値を問い直し、逆に学力の数値から教師の実践知を問い直すのが、DS-EFAである。

11. 数値から実践感覚を問い直す分析結果

DS-EFAの構想は、まだ途についたばかりである。それでも、これまでの試行錯誤の中から、わずかながら貴重な分析結果を得ている。以下、教師たちの実践感覚を数値によって問い直す結果になった分析結果を示しておきたい。

教師に限らずわたしたち大人は、感覚や経験値やテスト結果や先入観から、「この子は学力が高い、この子は中くらい、この子は低い」とみなしがちではないだろうか。また学力を固定的にとらえ、学力が「高どまりの子」とか「低迷したままの子」などとみなす傾向がありはしないだろうか。1年とか2年のスパンで見れば、こうした実践感覚にも妥当性があるかもしれない。ところが、8年という長期スパンで見ると、この実践感覚には大きな修正が迫られる。小学1年から中学2年までの8年間に、個々の子どもの「学力」は大きく変動する。以下はそれを実証するエビデンスである。

わたしたち学力・非認知能力調査研究チーム²¹は、宮城県のM自治体の協力を得て、東京書籍（以下「東

²¹ 科学研究費助成事業、基盤研究 (B) 「グローバル世界を視野とする学力・非認知能力の効果的学校モデル」(課題番号：20H01667、研究代表者：田端健人)

書)の学力調査個票を貸与いただいている。M自治体は総人口2万5千人ほどであり、3つの公立小学校、2つの公立中学校をもつ。M小学校の児童はほとんどがM中学校に進学する。M自治体は長年、東書学力テストを4月と12月に実施してきた。12月テストは、経年変化が測定できる設計になっているという²²。本調査研究では、同テストの全国偏差値を「学力」と操作的に定義した。偏差値を「学力」と定義する手法は、今日の高度な学力研究でも用いられており(cf., 赤林ほか, p.34, p.63: 日本財団, p.14)、一定の妥当性をもつ。全国偏差値を「学力」とみなすということは、全国の同一学年集団の中での個人の相対的な位置を「学力」とみなすことである。この定義では、学力が上がるということは、全国の同一学年集団の中での順位が上がることを意味する。

本調査研究では、M小学校の平成24年度入学児童の全国偏差値を、M中学校2年生まで8年間追跡した。これだけの長期追跡調査は、現時点の日本の学力研究では稀である。紙幅の都合上、ここでは、算数・数学の全国偏差値の経年変化のみを紹介する。転校とかテスト当日の欠席とかで追跡できなかった生徒を除く、113名のデータをもとにしている。

8年間の個人の偏差値の変動幅(最大値と最小値との差)を計算したところ、8年間で偏差値が10(1シグマ)以上変動した児童は全体の71.4%、偏差値が20(2シグマ)以上変動した児童も全体の14.3%いることがわかった。

個人の8年間の偏差値の平均をとり、その平均で最上位から最下位までを並べ、四分分割し、最上位群、上位群、下位群、最下位群に分け、各群ごとの偏差値の変動幅を箱ヒゲ図にしたのが、図8である。

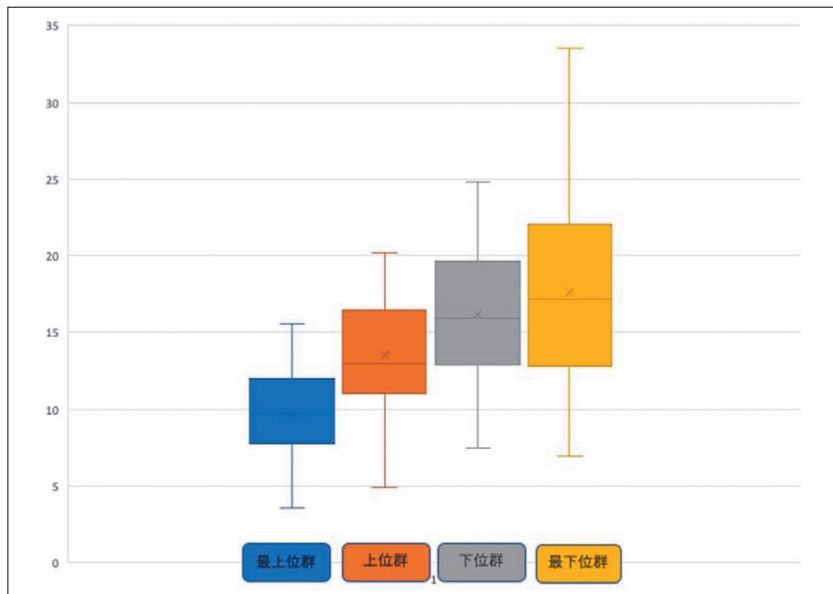


図8 M自治体113名の児童生徒の算数・数学の全国偏差値8年間の四分位ごとの変動幅

縦軸は偏差値の変動幅である。最上位群の変動幅の中央値が最も低いですが、それでも中央値で偏差値にしておよそ10動いている。群の順位が下がるにつれ、変動幅の中央値は上がってゆく。これは、学力が低い子どもほど、学力の変動幅が大きいことを意味している。この箱ヒゲ図は、偏差値が20以上変動する全体の

²² 同テストの特徴については、以下のURLで紹介されている。
https://www.tokyo-shoseki.co.jp/academic/n_hyoujun.html (最終閲覧日2020.11.20)

14.3%のほぼ全てが下位群と最下位群の子どもであることも可視化している。これは、学力下層の子どもの伸び代の大きさを示すと同時に、下がり幅の大きさも示している。最上位群が比較的安定的ではあるが、それでも中央値が10あたりであることからすれば、どの子どもの学力も、かなり可変的であるとみなしてよい。どの子どもも、本人、家庭、教師、仲間、学級、学校など様々な要因によって、年度年度の学力を良くも悪くも変動させるのである。わたしたち大人は、「この子は勉強ができる子」「この子は勉強ができない子」と決めつけてはならない。すべての子どもが学力を伸ばす可能性を秘めている。

またわたしたち大人の実践感覚には、「家庭環境が厳しい児童生徒の学力は低迷する」という先入観もある。先入観だけでなく、家庭環境、専門用語では「社会経済的地位 (Socio-Economic-Status : SES) と学力とが相関することを実証する研究も少なくない。

本調査研究チームは、上記の113名をよく知る教師の協力を得て、113名の中から「家庭の経済状況が厳しい」と感じられる子ども16名をピックアップしてもらった。これは、家庭訪問や面談などを通して教師たちが感覚的につかんだ実践知であり、主観的ではあるが、DS-EFAでは、本稿8節のように、このての教師の実践感覚は少なからず精度が高いと考える。少なくとも、家庭の蔵書数を児童生徒に質問し、それをSESの代替変数とするよりは、精度が高いと考えている。

この16名の8年間の偏差値平均を上記の四分位に分類すると、最上位群1名、上位群1名、下位群7名、最下位群7名という結果になった。わたしたちの調査研究でも、平均で見れば、やはりSESと学力には一定の相関が推測される結果となった。ただし、ここでは図示しないが、これら16名の偏差値変動幅の箱ヒゲ図を描くと、それは図8の「下位群」とほぼ同じ形になった。これも、従来の学力研究の知見「Lowest SESにおいて学力のばらつきが大きい」(お茶の水大学, p.14) と整合的である。この先行研究が、平成29年度全国学力・学習状況調査の結果であり、「ばらつき」はLow SES集計児童生徒集団の同一時点でのスコアの散らばりを意味するのに対して、わたしたち分析の「ばらつき」はLow SESと教師が判断した児童個人の8年間の経年の偏差値の変動である。

Low SESと判断される16名の中から、8年間で偏差値が20以上変動した児童5名の経年偏差値をグラフ化すると、図9になる。

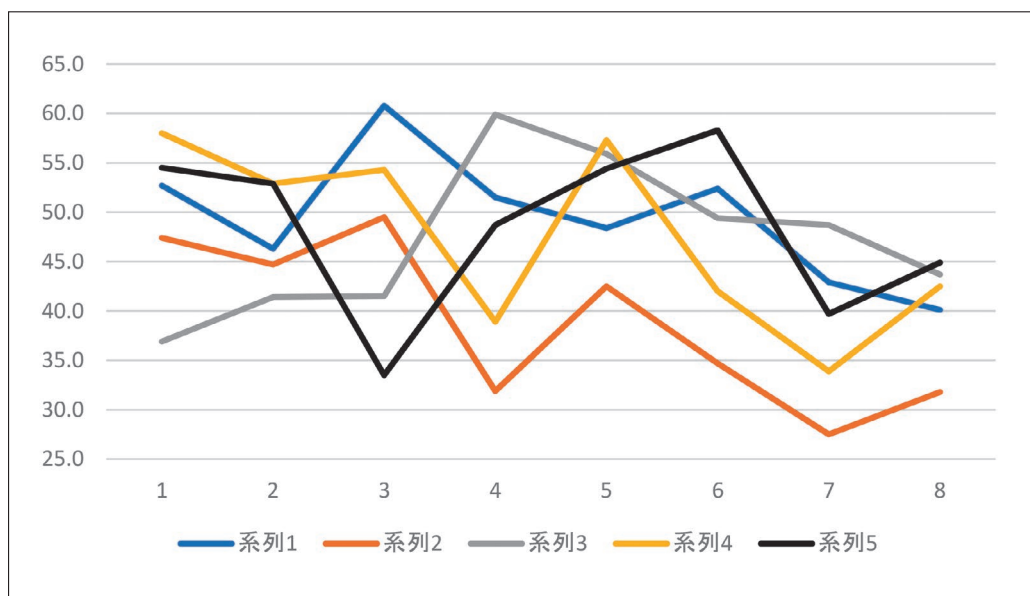


図9 Low SESで偏差値が20以上変動した児童生徒の偏差値経年変化

それぞれの折れ線(系列1~5)は個人の偏差値の経年変化を表している。縦軸が毎年のテストの偏差値、横軸が学年である。例えば、黒線(系列5)の子どもの8年間の平均偏差値は48.4であり、最高偏差値58.3、最低偏差値33.5であり、その差は24.8である。グラフを見てわかる通り、Low SESの子どもも、高い学力を示すことが十分あり、家庭環境だけで学力が決まるわけではない。子ども本人と教師たちと仲間集団が力を合わせれば、Low SESという不利な条件を克服する余地は十分にある。

12. なぜ「学力」にこだわるのか？

ここで改めて「教育生産関数」という考え方を思い出したい。「教育生産関数」とは、「学力」という教育の成果を、「〇〇をしたから、学力が上がった」といった単純な因果関係ではなく、多種多様な要因が複雑に作用し合う「関数」として捉えるコンセプトである。わたしたちは暫定的に、次のような教育生産関数を考えている²³。

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

表6 わたしたちが考える主な投入要素

X (生徒本人の気質や頭脳)	生まれつきの気質や体力や頭脳など。ただし、わたしたちは、頭脳が一生にわたり構造的質的に変化すると考えている。
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)	教育行政、学校・学級経営、地域と学校との関係、教師の人柄や技量など
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
N (非認知スキル)	自制心、自己効力感、勤勉性、協調性、やり抜く力、ポジティブ度など

添字の i は個人を、 t は時間を表している。

SESは確かに統計的に有意とされることの多い変数ではあるが、他の諸要因によって挽回できないわけではない。Low SESという困難な状況にある子どもにも、確かな学力を保障することは、公教育の一つの使命であろう。

またこの教育生産関数は、なぜわたしたち調査研究チームが「学力」にこだわるのか、を説明するためにも有効である。わたしたちが「学力」にこだわるのは、点取合戦に興味があるからではない。そうではなく、「学力」が子どものSESや教師の働きかけやクラスメイトたちとの関わり、そして自制心や自己効力感や協調性といった非認知スキルなどの総体の出力の一つだからである。子どもを成長させる多種多様な要因の作用の総体が、「学力」となって現れる、と考えるからである。一人ひとりのテスト結果の数値の背後に、上記の諸変数すべてを読み取り、阻害要因の改善を図ろうとするからである。

DS-EFAの構想は、すべての子どもの学力と非認知スキルの潜在力を信じ、どのような教育行政、学級・学校経営、授業づくり、学習法、行事や特別活動のあり方が、子どもの学力と非認知スキルを育てるかを、数値データや感覚データによって解明し、すべての子どもたちの現在と将来のWell-Beingを高めることを目指している。

²³ 代表的な投入要素については、石川ほか(2017), p.4を参照。

【謝辞1】 統計学、特に項目反応理論の第一人者である東北大学大学院教育学研究科・柴山直教授のご教示ご助言のおかげで、平均に関する本稿の着想を得ました。本稿3・4節のRによるK-S検定もご指導の賜物です。この場をお借りして、深く感謝申し上げます。もちろん、本稿の至らなさの責任は、ひとえに筆者である田端にあります。

【謝辞2】 文部科学省総合教育政策局調査企画課学力調査室にて、「全国学力・学習状況調査」の個票匿名データ貸与について対応くださったスタッフの皆様にも、この場をお借りして心から感謝申し上げます。データ貸与の手続きのプロセスで、全国学調やその調査分析が「序列化や競争」のためではないこと、むしろそれを抑制することを貴省が切望されていることを実感しました。本稿7節の仮説1は、貸与データを活用した今後の分析のための前提となります。

【謝辞3】 固有名を記すことはできませんが、11節で紹介した調査分析の土台となる個票データを貸与くださった教育委員会様にも、この場をお借りして深く感謝申し上げます。おかげで非常に貴重な知見が得られています。貴自治体はもちろん、全国の児童生徒のためになる分析を継続してまいります。

【付記】 本稿は科学研究費助成事業、基盤研究 (B) 「グローバル世界を視野とする学力・非認知能力の効果的学校モデル」(課題番号：20H01667、研究代表者：田端健人) の研究成果の一部である。

引用文献

- 赤林英夫・直井道夫・敷島千鶴編著, 2016『学力・心理・家庭環境の経済分析——全国小中学生の追跡調査から見てきたもの』有斐閣。
- Husserl, E. (1976a) *Die Krisis der europäischen Wissenschaften und die transzendente Phänomenologie*, Martinus Nijhoff. (邦訳) フッサール, E. (1974) 『ヨーロッパ諸学の危機と超越論的現象学』細谷恒夫・木田元訳, 中央公論社。
- 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.以下のURLからアクセス。
<https://www.rieti.go.jp/publications/pdp/17p010.pdf> (最終閲覧日2020.11.20)
- 石村貞夫, 2014『入門 はじめての統計解析』東京図書。
- 数実浩佑, 2019「Rを用いたPISA・TIMSSの分析」志水宏吉監修・川口俊朗編『日本と世界の学力格差』明石書店。
- 小島寛之, 2020『完全独習 統計学入門』ダイヤモンド社。
- メルロー＝ポンティ, M. 1989『知覚の現象学 1』竹内芳郎・小木貞孝訳, みすず書房。
- 日本財団, 2018「家庭の経済格差と子どもの認知能力・非認知能力格差の関係分析——2.5万人のビッグデータから見てきたもの——」(PDF) 以下のURLからアクセス。
https://www.nippon-foundation.or.jp/app/uploads/2019/01/wha_pro_end_07.pdf (最終閲覧日2020.11.20)
- お茶の水大学, 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」(PDF) 以下のURLからアクセス。
https://www.mext.go.jp/component/a_menu/education/micro_detail/_icsFiles/afeldfile/2018/07/10/1406896_1.pdf (最終閲覧日2020.11.20)
- 東北大学, 2014「東日本大震災の学力への影響——IRT推定値による経年比較分析——」(研究代表者 柴山直). 本調査研究には、以下のURLからもアクセスできる。https://www.mext.go.jp/a_menu/shotou/gakuryoku-chousa/1361052.htm (最終閲覧日2020.11.20)

資料：K-S検定_Rスクリプト

```
# 2020.12.6
#
# コルモゴロフ=スミルノフ検定 (2 標本)
# Kolmogorov-Smirnov test
# https://bellcurve.jp/statistics/glossary/1360.html
#
```

```

# 全国学力・学習状況調査の平均正答率をどう受けとめるべきか？
# ——「生きられた数値」による〈子ども教育データサイエンス〉の構想——
# How to Understand Means of National Scholastic Tests? :
# Ideas of 'Data Science of Education for All' based on Lived Numbers
# 田端 健人
# Taketo TABATA

# Table1 H31結果比較
#
#
# ソースデータ
# https://www.nier.go.jp/19chousakekkahoukoku/factsheet/19prefecture-City/
#
# 各都道府県別エクセルファイル
# (例:北海道) 調査結果概況 北海道－児童（公立）(PDF/66KB)
# https://www.nier.go.jp/19chousakekkahoukoku/factsheet/19prefecture-City/01_hokkaido/01p_19r.xlsx
# の度数分布表（正答数集計）の欄の相対度数（割合）を使った。
#

# 小6国語

zenkoku_p6j <- c (4.8,9.7,11.9,12.3,11.4,10.2,8.7,7.2,6.0,4.9,4.0,3.2,2.5,1.7,1.2)
akita_p6j <- c (10.8,16.8,15.2,14.1,11.0,8.8,6.9,4.7,4.0,2.6,2.1,1.4,0.8,0.6,0.2)
aichi_p6j <- c (4.5,9.6,12.5,13.2,12.0,10.7,9.0,7.5,6.1,4.8,3.9,2.9,2.0,0.9,0.5)

ks.test (zenkoku_p6j,akita_p6j) #全国分布と秋田分布の同等性
ks.test (zenkoku_p6j,aichi_p6j) #全国分布と愛知分布の同等性
ks.test (akita_p6j,aichi_p6j) #秋田分布と愛知分布の同等性

# 小6算数

zenkoku_p6m <- c (4.8,9.7,11.9,12.3,11.4,10.2,8.7,7.2,6.0,4.9,4.0,3.2,2.5,1.7,1.2)
ishikawa_p6m <- c (8.9,15.4,15.5,14.1,11.7,9.0,6.5,5.3,4.0,2.8,2.2,2.2,1.4,0.7,0.4)
hokkaido_p6m <- c (3.9,8.9,12.1,13.1,11.9,10.5,9.4,7.9,6.8,5.5,4.3,3.1,1.7,0.6,0.2)

ks.test (zenkoku_p6m,ishikawa_p6m) #全国分布と石川分布の同等性
ks.test (zenkoku_p6m,hokkaido_p6m) #全国分布と北海道分布の同等性
ks.test (ishikawa_p6m,hokkaido_p6m) #石川分布と北海道分布の同等性

# 中3国語

zenkoku_m3j <- c (17.8,20.5,17.1,13.2,9.9,7.2,5.3,3.8,2.7,1.7,0.7)
akita_m3j <- c (21.9,25.0,17.5,12.7,9.1,5.9,3.8,2.1,1.4,0.6,0.2)
okinawa_m3j <- c (12.6,17.1,16.3,13.7,11.7,8.9,7.5,5.1,3.8,2.5,0.9)

ks.test (zenkoku_m3j,akita_m3j) #全国分布と秋田分布の同等性
ks.test (zenkoku_m3j,okinawa_m3j) #全国分布と沖縄分布の同等性
ks.test (akita_m3j,okinawa_m3j) #秋田分布と沖縄分布の同等性

# 中3数学

zenkoku_m3m <- c (3.6,7.7,9.5,9.7,9.1,8.2,7.3,6.7,6.2,5.9,5.6,5.3,4.8,4.2,3.4,2.0,0.8)
fukui_m3m <- c (4.9,11.0,12.7,11.7,10.2,7.8,6.8,5.8,5.7,4.8,4.8,4.2,3.2,2.8,2.0,1.2,0.5)
okinawa_m3m <- c (2.8,5.3,6.9,7.3,7.2,6.7,6.7,6.6,7.0,7.4,7.0,6.6,6.4,6.4,5.1,3.4,1.1)

ks.test (zenkoku_m3m,fukui_m3m) #全国分布と福井分布の同等性
ks.test (zenkoku_m3m,okinawa_m3m) #全国分布と沖縄分布の同等性

```

```
ks.test (fukui_m3m,okinawa_m3m) #福井分布と沖縄分布の同等性

# Table2 H19結果比較
#
# ソースデータ
# https://www.nier.go.jp/tyousakekka/06todofuken\_chousakekka\_shiryoutm

# 中3 数学B 全国調査結果概要の正答数集計値（児童数）

# 0問～17問の全国生徒数を転記（相対度数のエクセルファイルがないため）
zenkoku_m3mn<-c (14445, 18511, 23789, 29147, 36799, 46338, 55515, 62444, 67279, 70008, 73080, 76874, 83488, 91447, 99715,
101568, 84389, 42373)

# 全国生徒数
# 1077209

# 全国の中3 数学Bの相対度数（割合）の算出
zenkoku_m3mp<-zenkoku_m3mn/1077209*100

# 少数第1位以下四捨五入
round (zenkoku_m3mp, digits=1)

# 結果（中3 数学Bの全国相対度数）を格納
zenkoku_m3m<-c (1.3, 1.7, 2.2, 2.7, 3.4, 4.3, 5.2, 5.8, 6.2, 6.5, 6.8, 7.1, 7.8, 8.5, 9.3, 9.4, 7.8, 3.9)

# 中3 数学B 福井の正答数集計値（児童数）
fukui_m3mn<-c (39, 60, 95, 126, 159, 245, 278, 334, 404, 497, 553, 602, 645, 703, 846, 914, 808, 448)

# 福井の生徒数
# 7756

# 福井の中3 数学Bの相対度数（割合）の算出
fukui_m3mp<-fukui_m3mn/7756*100

# 少数第1位以下四捨五入
round (fukui_m3mp, digits=1)

# 結果（中3 数学Bの福井の相対度数）を格納
fukui_m3m<-c (0.5, 0.8, 1.2, 1.6, 2.1, 3.2, 3.6, 4.3, 5.2, 6.4, 7.1, 7.8, 8.3, 9.1, 10.9, 11.8, 10.4, 5.8)

# 中3 数学B 沖縄の正答数集計値（児童数）
oki_m3mn<-c (383, 514, 717, 832, 1034, 1229, 1378, 1307, 1290, 1223, 1146, 974, 902, 891, 793, 613, 371, 152)

# 沖縄の生徒数
# 15749

# 沖縄の中3 数学Bの相対度数（割合）の算出
oki_m3mp<-oki_m3mn/15749*100

# 少数第1位以下四捨五入
round (oki_m3mp, digits=1)

# 結果（中3 数学Bの福井の相対度数）を格納
oki_m3m<-c (2.4, 3.3, 4.6, 5.3, 6.6, 7.8, 8.7, 8.3, 8.2, 7.8, 7.3, 6.2, 5.7, 5.7, 5.0, 3.9, 2.4, 1.0)
```

```
zenkoku_m3m<-c (1.3, 1.7, 2.2, 2.7, 3.4, 4.3, 5.2, 5.8, 6.2, 6.5, 6.8, 7.1, 7.8, 8.5, 9.3, 9.4, 7.8, 3.9)
fukui_m3m<-c (0.5, 0.8, 1.2, 1.6, 2.1, 3.2, 3.6, 4.3, 5.2, 6.4, 7.1, 7.8, 8.3, 9.1, 10.9, 11.8, 10.4, 5.8)
oki_m3m<-c (2.4, 3.3, 4.6, 5.3, 6.6, 7.8, 8.7, 8.3, 8.2, 7.8, 7.3, 6.2, 5.7, 5.7, 5.0, 3.9, 2.4, 1.0)
```

```
ks.test (zenkoku_m3m,fukui_m3m) #全国分布と福井分布の同等性
```

```
ks.test (zenkoku_m3m,oki_m3m) #全国分布と沖縄分布の同等性
```

```
ks.test (fukui_m3m,oki_m3m) #福井分布と沖縄分布の同等性
```

```
# end of R script
```