

## 国際学力調査からみる日本の学力の変化

### Trends in International Test Scores in Japan

川 口 俊 明

Toshiaki KAWAGUCHI

学校教育講座

(平成25年9月30日受理)

#### 1. はじめに

本稿の目的は、TIMSS・PISA という2つの国際学力調査をもとに、2000年以後の日本の学力の変化をまとめ、日本の現状と課題について整理することである。

2000年代初頭に始まった学力低下論争以後、TIMSS・PISAは、日本では、その結果が常に注目を集めてきた。とくにPISAは、「PISA型学力」といった言葉ができるほど、注目を集め、教育現場に大きな影響を与えている。2007年に開始された全国学力・学習状況調査にはPISAを意識したB問題が組み入れられ、学習指導要領にもPISAの「応用力」「活用力」に触れた記述が見られる。PISAに比べれば影響力は小さいものの、TIMSSの注目度も低くはなく、各回の結果が報告されるたびに、日本の順位に一喜一憂する教育関係者は多い。

こうした状況を見ても、国際学力調査が日本の教育に与えた影響はかなり大きいと言えるだろう。しかし、日本で行われる学力に関する議論を見ると、本当に国際学力調査の結果を理解しているのだろうか疑問に感じることも多い。たとえばPISAに対する関心の高さとは裏腹に、報告書の日本語訳すら十分に行われていない。PISA2009の報告書を見ると、日本語版(国立教育政策研究所2010)は300ページに満たない。これに対して、英語版(OECD2010)は200～300ページの6分冊構成だから、全部で1000ページを超えている。記述の重複があるとは言え、それを省いたとしても、日本語訳された部分があまりに少ない。TIMSSの報告書の日本語訳も、その一部が国立教育政策研究所のホームページで公開され、書籍化されたものは絶版という状況である。こうした事態は、日本でTIMSS・PISAという調査の名前と結果のごく一部だけが一人歩きし、調査内容が十分に省みられていない一つの証左であるように思う。

TIMSS・PISAは、2013年時点で、日本の教育研究者・政策立案者が参照できる、もっとも優れた学力調査である。その理由は、①日本全体を偏りなく調査対象にした学力調査であること、②二時点以上の得点の変化が測定されており、日本の学力の変化がわかること、③子どもたちの家庭背景に関する情報が含まれていること、④すべてのデータがウェブ上で公開されており、誰でも利用できること、の四点に求められる。これらの点について、少し詳しく説明しておこう。

2000年以降、学力低下・学力格差論に注目が集まり、教育社会学者をはじめとして、研究者による学力調査が実施されてきた(荻谷・志水2004, 耳塚2007)。たしかに、これらの学力調査は、子どもたちの学力低下・学力格差を指摘したという点で一定の意義は持っているものの、その多くが特定の地域や協力を得られた一部の学校を対象に実施されたものである。つまり、仮にその調査で成績の変化や格差の存在が明らかになったとしても、それが日本全体の傾向なのかどうかははっきりしない。これは、教育研究・教育政策を論じる上で非常に大きな問題である。学力低下・学力格差の実態がそれほど知られていなかった2000年前後であればともかく、一定程度その存在が認知された現在では、やはり日本全体の傾向がどうなのかということにもっと関心を向けなければならないだろう。

日本全体の傾向を知るという点では、2007年より実施されている全国学力・学習状況調査は一定の意義

を持っている。ただし、この調査は、大きな問題をいくつも抱えている<sup>(1)</sup>。その一つが、年度ごとの成績の変化を捉えることができないという点である。子どもたちの成績が変化したとして、それが調査の難易度に由来するものなのか、それとも子どもたちの能力が変化したためなのかという点を識別できないのでは、何のために学力調査を実施したのかわからない。この点、TIMSS・PISAは、項目反応理論（PISAはラッシュモデル）と呼ばれる測定理論に基づき、成績の変化を捉えることが可能な設計になっている<sup>(2)</sup>。

全国学力・学習状況調査には他にも大きな欠点がある。それは、子どもたちの家庭環境をほとんど把握していないという点である。一般に、家庭環境の恵まれた子どもほど成績が高くなりがちである。このことを踏まえるなら、子どもの成績には、政策の影響や学校・教師の努力だけでなく、家庭環境の要因も少なからず影響していることになる。そのため、学力調査を分析する際には、家庭環境が成績に与えている影響をできるだけ考慮しながら、政策の影響、学校・教師の努力を把握しなければならない。TIMSS・PISAは、両親の学歴や年収を含めた、家庭環境に関する質問を実施しており、その点でも非常に有益な調査である。

さらに、TIMSS・PISAは、その設計もさることながら、データがウェブ上ですべて公開されている<sup>(3)</sup>。そのため、インターネットに接続できる環境にある人なら、誰でもこの調査を再分析することができる。学力調査に関心のある研究者はもちろん、行政の担当者や、一般の人々でも、誰でも調査を再分析し、データに基づいて議論を行うことができる。誤った議論や分析であれば、他の誰かが指摘し訂正ができるという点も、TIMSS・PISAの優れた点である。

残念ながら、2013年時点で日本には、上記の点を同時に満たす学力調査はTIMSS・PISA以外には存在していない。こうした点を踏まえれば、学力について議論をしたり、教育政策を考えたりする上で、2つの国際学力調査の概要を把握しておくことは重要である。

以下では、2つの国際学力調査をもとに、日本の学力の変化と課題について見ていくことにしたい。その際に注目するのは、次の2点である。1つは、学力は低下したのか否かという点である。学力低下論争以後、日本の子どもたちの学力が低下したのかどうかは、常に学力をめぐる論争の中心にあった。本稿では、国際学力調査の日本の成績に焦点を絞り、学力が低下したのか否か確認してみたい。

もう1つは、学力の格差は拡大したのか否かという点である。学力低下論争で話題になったことの一つに、成績の低下のみならず、家庭環境のちがいによって生じた学力格差が存在し、かつそれが拡大しているのではないかといった学力格差拡大の議論があった（荻谷・志水2004）。こうした議論は、おもに教育社会学者を中心に提出されてきたが、すでに触れたように、その多くは偏った学力調査から結論を導いている。国際学力調査をもとに、あらためて現在の日本の状況について整理することが、本稿の目的である。なお、同様の問題意識を持ったものとして、須藤（2013）があるが、これは2007年までの国際学力調査をもとに、中学校の理数系学力に限定したものである。本稿では、国際学力調査の全領域を対象に、学力低下・学力格差について議論を行う。

## 2. 国際学力調査の概要

得点を見る前に、TIMSS・PISAの概要と、調査結果を読みとる際に知っておかなければならない基本的な情報について整理しておこう。はじめに、TIMSSの概要を、続いてPISAの概要について述べていく。

### 2.1. TIMSSの概要

TIMSS (Trends in International Mathematics and Science Study) は、IEA (The International Association for the Evaluation of Educational Achievement: 国際教育到達度評価学会) が、おもに各国の4年生・8年生を対象として実施している算数・数学及び理科の到達度に関する国際学力調査である。日本は1995年の第1回TIMSSから、以降のすべての調査に参加している。TIMSSで出題されている問題は、日本の小中学校の試験で出題される算数・理科の問題と似通っており、いわゆる算数のテスト、理科のテストと考えても、それほど間違いではない。本稿では、問題内容について検討する紙面の余裕はないが、国立教育政策研究所のホームページ等に日本語版の問題文が掲載されているので、問題に興味のある方は、そちらを参照されたい。

TIMSSの結果を見るときには、いくつか注意が必要な点があるが、とくにTIMSS1995とTIMSS1999は、TIMSS2003以降と多少性格が異なっているという点は重要である。TIMSS1995では、①調査時に9歳の子どもを多く含む2つの学年、②調査時に13歳の子どもを多く含む2つの学年、③中学校の最終段階にいる

子ども、という3つの集団が、調査対象となっていた。日本は③の調査には参加していないが、前2者に参加している。つまり、TIMSS1995には、日本から小学3年生・4年生と中学1年生・2年生の4学年が参加していたのである。4年後のTIMSS1999は調査対象を限定し、13歳の子どもを含む2つの学年のうち、より多くの13歳を含む学年を対象にしている。これは、日本の場合、中学2年生が該当する。調査対象の変更に伴い、TIMSS1999では、小学4年生は参加していない。

TIMSS2003以降では、調査対象が9歳を多く含む学年、および13歳を多く含む学年に変更された。これは日本では小学4年生および中学2年生に該当する。以降のTIMSSでは、この2つの学年が調査対象となっている。要するに、TIMSSの調査対象は、日本では基本的に小学4年生と中学2年生だが、TIMSS1995ではこれに加えて小学3年生と中学1年生のデータが利用できる。また、TIMSS1999は中学2年生のデータのみである。

なお、TIMSSの得点は、1995年の参加国の平均が500、標準偏差が100になるように調整されており、1999年以降の各国の成績の変化が分析できるようになっている。ただし、TIMSS1995の報告書だけは、後のTIMSSと異なる方式で計算された得点が表記されている。これは、TIMSS1999で、以後のTIMSSと比較可能にするために、TIMSS1995の得点を再計算したためである。本稿のTIMSS1995の得点は、このTIMSS1999で再計算された数値を使用しているため、TIMSS1995の報告書に記載された値とは異なっている箇所がある。こうしたTIMSSの調査設計については、Martin et al (2012) 等に詳しい。

## 2.2. PISAの概要

PISAは、OECD (Organisation for Economic Co-operation and Development: 経済協力開発機構) によって2000年から3年おきに実施されている国際学力調査である。その目的は、各国の子どもたちが将来生活していく上で必要とされる知識や技能が、義務教育修了段階である15歳時点において、どの程度身につけているかを測定することであり、日本では高校1年生を対象に7月に調査が実施されている。PISAでは、読解力 (Reading Literacy)、数学的リテラシー (Mathematics Literacy)、科学的リテラシー (Science Literacy) を主要な調査領域としている。

PISAの結果を解釈するときに注意しなければならない点は、PISAの主要な調査領域で出題されている問題が、私たちが想像する国語・算数・理科といった教科とは大きく異なっているという点である。とくに、PISAの読解力で測定されている能力は、日本で言う国語とはかなり異なっている。PISAの読解力の点数を高めるために読書を奨励する人がいるが、これは誤っているとは言わないまでも、あまり適切な対策ではない。PISAの調査問題については、国立教育政策研究所のホームページ等に日本語版の問題文が掲載されているので、具体的な問題文については、そちらを参照されたい。また、PISAの問題文については、さまざまな研究者が検討を重ねているが、村山 (2006) や全国大学国語教育学会編 (2010) は、問題文に即した検討が行われており、PISAを理解する上で参考になる。

PISAでは、調査時間の2/3の時間を費やして主要分野 (Main domain) を重点的に調べ、残りの領域 (Sub domain) については概括的な状況を調べることになっている。主要分野は、2000年と2009年が読解力、2003年と2012年が数学的リテラシー、2006年と2015年が科学的リテラシーである。すべての生徒は、主要分野のなんらかの問題を解くが、それ以外の分野については、調査設計の関係上、まったく問題を解かない生徒もいる。こうした事情があるため、細かい分析や多変量解析等の統計手法を使用した分析は、主要分野で行われることが多い。

各領域の得点は、はじめて主要分野となったときのOECD加盟国の平均が500、標準偏差が100になるように調整されている。たとえば数学的リテラシーの得点は、2003年のOECD加盟国の平均が500・標準偏差が100になるように調整されている。2003年以降の数学的リテラシーの得点は、2003年の得点を基準に計算されており、相互に比較することが可能である。

ここで一つ注意しなければならないことがある。それは、PISAの得点は経年比較が可能だが、それはその領域が主要分野となった後に限られる。つまり、2000年と2003年の読解力や、2006年と2009年の数学的リテラシーの得点は比較可能だが、2000年と2003年の数学的リテラシーや、2003年と2006年の科学的リテラシーの得点を比較することはできない。PISAの結果を見るときは、比較可能な箇所と比較できない箇所を区別して議論する必要がある。こうしたPISA調査の詳細については、PISAのTechnical Report (OECD 2012) 等を参照してほしい。

### 3. 日本の学力の変化

以上を踏まえた上で、TIMSS・PISA から何が言えるのかについて見ていくことにしよう。はじめに、学力は低下したのか否かについて議論した後、学力の格差について議論を進めることとする。

#### 3.1. 学力は低下しているのか？

日本の子どもたちの学力が低下したかどうかという問題は、2000 年以降、常に学力をめぐる論争の中心的な話題となってきた。しかし、冒頭でも述べたように、日本全体の学力の変化を示す証拠となるような学力調査は、ほとんど存在していない。そんな中、数学と理科については、TIMSS がかなり説得力のある根拠を提示してくれている。

表 1. TIMSS 小 4 算数の変化

		平均	標準誤差	1995	2003	2007
1995	小 3	512.8	1.7			
1995	小 4	567.2	1.9			
2003	小 4	564.6	1.6			
2007	小 4	568.2	2.1			
2011	小 4	585.4	1.7	↑	↑	↑

表 2. TIMSS 小 4 理科の変化

		平均	標準誤差	1995	2003	2007
1995	小 3	501.0	1.6			
1995	小 4	553.2	1.7			
2003	小 4	543.5	1.5			
2007	小 4	547.8	2.1			
2011	小 4	558.6	1.9	↑	↑	↑

はじめに、表 1・表 2 を見てみよう。これは、小学 4 年生の算数・理科の成績と、その標準誤差である。なお、すでに述べたように、TIMSS1995 は小学 3 年生・中学 2 年生も調査対象としているため、その成績も併せて記載している。

成績の変化について考えるときに重要なのは、成績には誤差があるという点である。TIMSS に限らず、多くの学力調査では、一部の子どもをサンプリングする標本調査が用いられている。そのため、単なる平均点のみならず、誤差を同時に報告することが一般的である。誤差を報告せず、上がった／下がったという議論をしている場合が見られるが、それは誤りである。誤差をもとに成績の変化を議論する場合、一般には、5% 水準で有意な差があるかどうかを問題にすることが多い。二時点（a と b）の成績の変化が有意であるか否かは、下記の式で求めることができる。なお、M は平均値、SE は標準誤差である。

$$\frac{M_a - M_b}{\sqrt{SE_a^2 + SE_b^2}}$$

たとえば、1995 年の小学 4 年生の成績（平均 567.2、標準誤差 1.9）と、2011 年の成績（平均 585.4、標準誤差 1.7）の変化について考えてみよう。上の式に当てはめれば、次のようになる。

$$\frac{567.2 - 585.4}{\sqrt{1.9^2 + 1.7^2}} = \frac{-18.2}{2.55} = -7.14$$

ここで、計算された数値の絶対値が 1.96 を越えていれば、5% 水準で有意な差があると判断する。この場合は、有意な差が存在するため、1995 年度と 2011 年度で、小学 4 年生の算数の成績は上昇したと判断することになる。こうした処理を行い、成績の変化を示したのが表 1 や表 2 の右側である。矢印のついている箇

所が成績が上昇した、あるいは下降したと判断できる箇所である。成績の変化が確認できない場合は、空欄になっている。

表1を見ると、算数に関しては、日本の小学4年生の成績はむしろ向上している。とくに、2011年の成績は、1995年時点よりも高く、学力低下とは言えない状況になっている。表2の理科については、2003および2007の成績が1995年と比べて下回るものの2011の成績は、過去のすべてのTIMSSの成績を上回っており、こちらも学力が低下したとは言い難い結果である。

こうした結果が生じた理由については想像するしかないが、一つは学力低下論争や全国学力・学習状況調査等が、学校現場の教員たちを学力向上に煽ったという側面はあるかもしれない。なお、新学習指導要領が実施されたのは、小学校でも2011年だから、学習指導要領の影響と言うには結果が出るのが早すぎるだろう。

次に、中学校の結果について見ていこう。表3が数学、表4が理科の結果を示したものである。変化を見る場合の手続きは、小学校と同じである。

表3. TIMSS 中2 数学の変化

		平均	標準誤差	1995	1999	2003	2007
1995	中1	556.2	1.7				
1995	中2	581.1	1.6				
1999	中2	578.6	1.7				
2003	中2	569.9	2.1	↓	↓		
2007	中2	569.8	2.4	↓	↓		
2011	中2	569.8	2.6	↓	↓		

表4. TIMSS 中2 理科の変化

		平均	標準誤差	1995	1999	2003	2007
1995	中1	521.1	1.8				
1995	中2	554.5	1.8				
1999	中2	549.7	2.1				
2003	中2	552.2	1.9				
2007	中2	553.8	1.8				
2011	中2	557.7	2.4		↑		

表3を見ると、中学2年生の数学の成績は低下していると言えるようである。とくに、2003年以降の成績は、1995および1999の成績より低い。中学2年生の数学に関しては、学力低下と言えそうである。ただし、低下したとはいえ、TIMSS1995の中学1年生の成績よりは高いため、そこまで大きな変化とは言えないだろう。理科については、数学とはちがひ、低下傾向は見られない。ほぼ横ばいといった解釈が妥当だと思われる。

次に、PISAの経年変化についてみていこう。すでに述べたように、PISAは、日本ではあまりなじみのない能力を測定している。そのため、PISAの成績が変化したからと言って、それを国語の成績が低下した、あるいは算数の成績が低下した・・・といった具合に読みとるのは誤りである。PISAの成績の変化が、具体的に、どのような能力の変化を意味しているのかという点については、慎重に議論しなければならない。

また、TIMSSと異なり、PISAで経年変化を判断する場合は、Link Errorと呼ばれる数値を考慮しなければならない。PISAやTIMSSでは、経年変化を可能にするために、あらかじめ2つの調査間で同一の問題を一部出題し、この問題への反応を見ながら、2つの調査間の得点を調整している。こうして計算された得点には、それぞれの調査の誤差だけでなく、2つの調査を結合した際の誤差が混ざることになる。これがLink Errorである。PISAのLink Errorは、表5のとおりである。すでに述べたように、PISAには主要分野になった時期のちがひにより、比較できない箇所が存在する。表5に記載が存在しない箇所（たとえば、数学リテラシーの2000→2003など）は、比較することができない。

表 5. PISA Link Error

読解力		数学的リテラシー		科学的リテラシー	
2000 → 2009	4.937	2003 → 2009	1.990	2006 → 2009	2.566
2003 → 2009	4.088	2006 → 2009	1.333	2003 → 2006	4.963
2006 → 2009	4.069	2003 → 2006	1.382		
2003 → 2006	5.307				
2000 → 2006	4.976				
2000 → 2003	4.474				

Link Error を使った比較は、下記の式で行われる。

$$\frac{M_a - M_b}{\sqrt{SE_a^2 + SE_b^2 + \text{Link Error}^2}}$$

先ほどの TIMSS の式と比べてみると、Link Error を考慮する分、分母が大きくなりがちである。そのため、PISA で学力が変化したかどうかの判断基準は、TIMSS よりやや厳しめに行われていると言えるだろう。なお、理論的には、TIMSS にも Link Error が存在するが、TIMSS では Link Error は報告されていない。

表 6. PISA 読解力の変化

	平均	標準誤差	2000	2003	2006
2000	522.2	5.2			
2003	498.1	3.9	↓		
2006	498.0	3.6	↓	↓	
2009	519.9	3.5		↑	↑

それでは、PISA の経年変化について、読解力から見ていこう。表 6 は、各回の読解力の得点を示したものである。これを見ると、日本の読解力の成績は、2003 年と 2006 年で低下し、2009 年で 2000 年と同水準まで戻ってきている。いったん低下し、後に回復したと言うふうに読める。

ただし、こうした読解力の変化に対する解釈には、慎重である必要がある。たとえば、PISA 設計者の一人である Wu (2009) によると、経年変化を可能にするための調整に使用する問題の性質によっては、得点が適切に計算されないことがあるという。Wu は日本の読解力を例に挙げ、PISA2000 の読解力で出題された問題の性質上、日本の PISA2003 および 2006 の読解力の得点が下方修正されやすいことを指摘している。つまり、2000 年から 2003・2006 年の日本の読解力の得点の変化は、日本の子どもたちの学力に変化があったわけではなく、たんに PISA の調査設計に伴う問題だった可能性がある。PISA では主要分野以外の大きな調査枠組みの変更が行われないため、2009 年にふたたび読解力が主要分野となるまで、Wu の指摘した問題は修正されていない。つまり、PISA2009 で日本の読解力の成績が上昇しているが、これは、子どもたちの能力が上昇したのではなく、単に Wu が述べた問題が、PISA2009 で解決されたためかもしれない。いずれが正しいのかはわからないが、Wu の指摘を踏まえるのであれば、この間の日本の読解力の成績の変化は、あまり深刻に捉えすぎない方がよい。

表 7. PISA 数学リテラシー／科学リテラシーの変化 (左：数学 右：科学)

	平均	標準誤差	2003	2006		平均	標準誤差	2006
2003	534.1	4.0			2006	531.4	3.4	
2006	523.1	3.3	↓		2009	539.4	3.4	
2009	529.0	3.3						



続いて、数学的リテラシーと科学的リテラシーについてみてみよう。こちらは、主要領域となった時期が遅いため、経年比較が可能な箇所は多くない。比較可能な箇所について見てみると、低下傾向・上昇傾向と呼べるほど、はっきりした傾向は出ていない。大幅な変化はないと考えるべきであろう。

ここまで、TIMSS・PISAの成績の経年変化を見てきた。日本の子どもたちの成績は、TIMSSの中学2年生の数学でやや学力低下の傾向が見られる。他方で、TIMSSの小学4年生については、算数・理科ともに学力が向上している。また、PISAには日本の子どもたちの成績が変化しと断定できるだけの根拠は乏しい。要するに、日本の子どもたちの成績は、下がっている箇所もあるが、上がっている箇所もあるし、変化がはっきりしない箇所も多いということである。少なくとも言えるのは、国際学力調査は、学力低下を強く主張する根拠にはなり得ないということであろう。

### 3.2. 格差は拡大しているのか？

子どもたちの属性（性別、家庭環境など）によって、学力に差が生じるのではないかという学力格差の話題は、学力低下論争以後、活発に議論されてきた。こうした論争で議論の中心となったのは、格差があるということだけでなく、その格差が拡大しているという点であった（荻谷・志水 2004）。ただし、日本全体を対象とした学力調査を欠くため、格差拡大を主張する議論に限界があることも事実である。ここでは、TIMSS・PISAをもとに、格差が拡大したかどうかについてみていくことにしよう。

なお、国際学力調査をもとに成績の変化を見るとき、そこには大きな壁がある。それは、日本の教育界では、子どもの属性（たとえば、性別、両親の学歴・年収・職業、家族構成、エスニシティなど）に由来する格差を議論することはタブー視されていた（あるいは、今でもタブー視されている）という点である。2013年度の全国学力・学習状況調査に、家庭環境に関する質問が盛り込まれたものの、それ以前は、こうした問題について議論するのは、教育社会学あるいは教育経済学に近い領域の研究者がほとんどであった。現在でも、研究者が学力格差に関する調査を行おうにも、学校あるいは行政に断られるという状況は、日本ではそれほど珍しくない。こうした事情もあって、学力格差問題が注目を浴びる前、すなわち2000年前後のTIMSS・PISAでは、日本の家庭環境に関する変数がほとんど欠損している。

さらに、変化を測定するには、子どもが回答した質問紙の内容が共通の質問項目でなければならないが、TIMSS・PISAともに、各回の調査で生徒質問紙の内容が変更されることも多いため、格差が拡大しているかどうかを判断するための指標となる質問は、それほど多くない。今回は、ほぼすべての調査で子どもたちが回答している「性別」と、1999年以降のTIMSSで毎回質問されており、かつ日本の回答がほとんど欠損していない「家庭の本の冊数」という変数をもとに、学力格差の変化について検討していく。

#### ①男女差

一般に、子どもたちの成績に男女差があることはよく知られている。古くは女性が学校へ行かなかった、あるいは行けなかったこと等を要因とした、女子の低学力が問題視されてきたし、現在では、逆に男子の低学力が問題視される国もある（Bunchmann et al 2008）。そんな中、日本では、成績の男女差が問題視されることは少ない。一つには、先に述べたような属性による差をタブー視する文化も影響しているかもしれない。しかし、TIMSS・PISAを含め、多くの国際的な学力調査の報告書には、成績の男女差に関する分析が設けられている。こうした国際的な傾向を踏まえるなら、日本でも男女差に対する注目は必要であろう。以下では、TIMSS・PISAをもとに、成績の男女差について見ておこう。

はじめに、TIMSSの小学4年生の成績を見てみよう。

表 8. TIMSS 小4 算数の男女差

		女子		男子		男女の得点差
		平均	標準誤差	平均	標準誤差	
1995	小3	510.9	1.9	514.7	2.2	-3.8
1995	小4	563.3	2.1	571.1	2.5	-7.8
2003	小4	562.7	1.8	566.3	2.0	-3.6
2007	小4	567.9	2.5	568.4	2.7	-0.5
2011	小4	583.9	2.0	586.8	2.4	-3.0

表 9. TIMSS 小 4 理科の男女差

		女子		男子		男女の得点差
		平均	標準誤差	平均	標準誤差	
1995	小 3	499.8	2.0	502.2	2.2	-2.4
1995	小 4	547.2	2.0	559.1	2.1	-11.9
2003	小 4	541.8	1.9	545.1	2.0	-3.3
2007	小 4	548.3	2.5	547.3	2.5	0.9
2011	小 4	555.9	2.6	561.3	2.0	-5.4

表 8 が算数の男女差，表 9 が理科の男女差を示したものである。わずかに男子の平均値が高いものの，この差はほとんどが統計的に有意な差ではない。小学校に関していえば，成績の男女差と，その変化はほとんどないと言ってよいだろう。続いて，中学校について見てみよう。

表 10. TIMSS 中 2 数学の男女差

		女子		男子		男女の得点差
		平均	標準誤差	平均	標準誤差	
1995	中 1	550.5	1.9	561.5	2.4	-11.0
1995	中 2	576.7	1.9	585.2	2.1	-8.5
1999	中 2	574.7	2.4	582.4	2.3	-7.7
2003	中 2	568.6	4.1	571.2	3.6	-2.6
2007	中 2	567.7	3.3	571.9	3.3	-4.2
2011	中 2	565.9	3.1	573.6	3.6	-7.7

表 11. TIMSS 中 2 理科の男女差

		女子		男子		有意差
		平均	標準誤差	平均	標準誤差	
1995	中 1	515.6	1.9	526.2	2.5	-10.7
1995	中 2	544.4	2.0	563.9	2.4	-19.5
1999	中 2	542.8	2.7	556.4	3.6	-13.6
2003	中 2	547.6	3.0	556.7	2.9	-9.1
2007	中 2	551.8	2.8	555.8	2.5	-4.0
2011	中 2	553.7	2.8	561.7	2.9	-8.0

小学校と異なり，中学校については，成績の男女差がはっきりと表れている。とくに，理科でその傾向が顕著であり，男子の方が女子より高い成績をとりやすい傾向が見られる。ただし，近年その差は縮まってきているようである。また，数学については，過去の TIMSS では男子の方が高かったが，最近の TIMSS では，そうした差は見られない。

次に，PISA について見てみよう。

表 12. PISA 読解力の男女差

	女子		男子		男女の得点差
	平均	標準誤差	平均	標準誤差	
2000	536.9	5.4	507.3	6.7	29.7
2003	509.0	4.1	486.6	5.5	22.4
2006	513.3	5.2	482.7	5.4	30.6
2009	539.9	3.7	501.0	5.6	38.9



表 13. PISA 数学リテラシーの男女差

	女子		男子		男女の得点差
	平均	標準誤差	平均	標準誤差	
2003	530.1	4.0	538.5	5.8	-8.4
2006	513.2	4.9	532.9	4.8	-19.7
2009	524.1	3.9	533.6	5.3	-9.4

表 14. PISA 科学リテラシーの男女差

	女子		男子		男女の得点差
	平均	標準誤差	平均	標準誤差	
2006	529.8	5.1	533.0	4.9	-3.3
2009	545.4	3.9	533.8	5.5	11.7

PISA 調査の場合、読解力の男女差が顕著である。すべての PISA で、女子が男子を大幅に上回っている。数学・科学リテラシーについては、数学的リテラシーで男子が女子を上回る傾向があるものの、2006 年を除いて、この差は統計的に有意ではない。TIMSS で見られる理科や数学の男女差が PISA で見られないのは、TIMSS と PISA が測定する能力の性質のちがいも影響していると思われる。

ここでは、男女差について見てきた。もっとも差が顕著なのは PISA の読解力である。ここでは、女子が男子を大幅に上回る傾向が続いていると言えるだろう。次に、TIMSS の中学 2 年生の理科および数学では、男子が女子を上回る傾向がある。ただし、その差は縮まりつつあり、近年の数学では差が見られなくなっている。小学校の TIMSS では、男女差はほとんど見られない。全体としてみれば、読解力を除き、日本の子どもたちの成績の男女差は小さく、その差も縮小傾向にあると言えるだろう。

## ②家庭にある本の冊数

近年の社会の変化や、教育改革に伴い、子どもたちのあいだにある学力格差が拡大しているのではないかという懸念は、教育社会学者を中心に、しばしば提出されてきたものである（荻谷 2008、恒吉 2008）。国際学力調査から、その状況についてみてみることにしよう。

表 15 および表 16 は、TIMSS の小学 4 年生の算数と理科の成績の変化を示したものである。まず、家庭にある本の冊数と成績は、強い関連を示しており、家庭にある本の冊数が多い子どもほど成績が高い。家庭にある本の冊数は、家庭の文化的経済的な豊かさや親の教育熱心さ等を示していると思われるが、こうした家庭的に恵まれた子どもほど、成績が高い傾向があるということである。

表 15. TIMSS 小 4 算数 家庭にある本の冊数

	0 - 10 冊	11 - 25 冊	26 - 100 冊	101 - 200 冊	200 冊以上	得点差
2003 年	527.9	549.4	574.7	587.1	590.1	62.2 (6.6)
2007 年	522.0	555.7	578.6	602.9	599.3	77.3 (7.3)
2011 年	540.6	572.0	594.9	612.6	615.4	74.8 (5.5)

( ) 内は標準誤差

表 16. TIMSS 小 4 理科 家庭にある本の冊数

	0 - 10 冊	11 - 25 冊	26 - 100 冊	101 - 200 冊	200 冊以上	得点差
2003 年	513.8	529.3	553.0	561.3	566.3	52.5 (6.5)
2007 年	511.0	538.1	555.6	574.2	576.9	65.9 (7.7)
2011 年	519.7	545.8	566.6	584.5	585.9	66.2 (6.1)

( ) 内は標準誤差

次に、表の右端の得点差の列は、家庭にある本の冊数が「0 - 10 冊」と「200 冊以上」の子どもの平均値の差を示したものである。これを見ると、算数・理科ともに、2003 年と 2007 年のあいだで、成績の差が大

きくなっているようである。算数では、2003年に62.6点だった得点差が、2011年には74.8点になっているから、12.2点の差が開いたことになる。理科も13.7点の差が開いている。ただし、括弧内を見ればわかるように得点差の標準誤差も大きいので、この間の変化は統計的には有意でない。小学4年生の数学と理科の学力格差は拡大傾向にあるものの、強く断定することはできないといったところであろう。

表 17. TIMSS 中2 数学 家庭にある本の冊数

	0－10冊	11－25冊	26－100冊	101－200冊	200冊以上	得点差
1999年	549.3	565.2	576.9	598.3	604.9	55.6 (6.7)
2003年	533.2	552.8	570.6	587.2	604.1	70.9 (4.3)
2007年	525.7	551.4	577.2	588.2	604.1	78.4 (6.3)
2011年	520.6	554.7	576.9	590.7	597.6	77.0 (6.5)

( ) 内は標準誤差

表 18. TIMSS 中2 理科 家庭にある本の冊数

	0－10冊	11－25冊	26－100冊	101－200冊	200冊以上	得点差
1999年	518.0	541.0	547.6	566.9	576.8	58.8 (6.1)
2003年	517.3	539.4	551.8	567.0	584.1	66.8 (4.2)
2007年	507.2	539.4	560.6	572.9	586.0	78.8 (6.0)
2011年	512.2	542.3	564.7	574.9	590.3	78.1 (5.2)

( ) 内は標準誤差

続いて中学2年生のデータを見てみよう。表17および表18を見ると、中学校でも得点差が拡大してきていることがわかる。1999年から2011年までのあいだに、数学は55.6点から77.0点と21.4点の差が開き、理科でも58.8点から78.1点と19.3点の差がついている。細かく見ると、数学では「200冊以上」の子どもたちの成績が横ばいであるのに対し、「0－10冊」の子どもたちの平均値が下がっているため、家庭環境的に恵まれない子どもたちの成績が低下して差が広がっていると考えられる。理科については、「0－10冊」の子どもたちの成績は横ばいかやや低下傾向であるのに対し、「200冊以上」の子どもたちの成績が向上傾向なので、家庭環境的に恵まれた子どもたちの成績が上昇した結果、差が広がっていると言えるだろう。

小学校よりも得点の変化が大きい分、中学校の得点差の変化は統計的にも有意であり、1999年と比較すると、2007年・2011年の調査では、数学・理科ともに、得点差が大きくなっていると判断できる。家庭環境による学力格差は、この10年間で拡大していると言ってよいだろう。

ここまで、本の冊数という限定的な指標ではあるものの、学力格差の変化について見てきた。平均値のみの単純な議論だが、学力格差が拡大傾向にあると言わざるを得ないだろう。とくに、その傾向は中学校で顕著であり、数学・理科ともに、格差が拡大傾向にある。

#### 4. まとめ

本稿では、TIMSSとPISAという国際学力調査を用い、日本の学力の変化について実態の把握と課題を明らかにしてきた。

得られた結果は次のとおりである。まず、経年変化については、TIMSSの中学2年生の数学を除けば、TIMSS・PISAともに、大きな平均値の変化は見られないか、あるいはむしろ向上している領域もある。少なくとも、算数・数学および理科の成績に関して言えば、過度に学力低下を心配する必要はないだろう。ただし、中学2年生の数学は、若干低下傾向なので、この点については、なんらかの対策が必要である。

次に、学力格差については、男女差と本の冊数による分析を行った。男女差については、PISAの読解力で女子が男子を大きく上回る状況が続いているが、それ以外の領域ではそもそも差が小さいか、あるいは差が縮小する傾向がある。他方で、本の冊数による成績の差を見た場合、子どもたちのあいだにある学力格差が拡大傾向にあることが見てとれる。とくに中学2年生の数学・理科でその傾向が顕著である。

まとめてみると、単純な学力低下論はとりあえず否定されたことになる。成績が低下している領域もたしかにあるが、成績がむしろ向上している領域もある。学年や教科のちがいのちがいの学力低下を

叫ぶ必要性は薄いだらう。他方で、学力格差については、あまり楽観できるような状況ではなさそうである。とくに、中学校の数学と理科で学力格差が拡大している傾向があり、この部分に、今後集中したてこ入れが求められている。

なお、こうした分析は、本来は国際学力調査ではなく、全国学力・学習状況調査をもとに行われるべきものである。2013年の全国学力・学習状況調査では、追加調査として、保護者質問紙や経年変化が加えられたが、それぞれの調査が別々のグループによって実施されるなど、非常に効率が悪い設計になっている。調査データの公開も含め、調査設計を抜本的に見直すべきであろう。

#### <注>

- (1)全国学力・学習状況調査の問題点については、荒居・倉元編（2008）、川口（2011）、木村（2006）、全国大学国語教育学会編（2010）などを参照されたい。
- (2)項目反応理論やラッシュモデルについては、豊田（2002）、静（2007）を参照。
- (3)TIMSS・PISAのデータおよび設計に関する情報は、次のホームページから入手できる。TIMSSについては、TIMSS & PIRLS (<http://timssandpirls.bc.edu>)。PISAは、My PISA (<https://mypisa.acer.edu.au>)。

#### <参考文献>

- 荒居克弘・倉元直樹編，2008，『全国学力調査日米比較研究』金子書房。
- Buchmann, C., DiPrete, T. & McDaniel, A., 2008, "Gender Inequality in Education," *Annual Review of Sociology*, 34(1), 319-337.
- 荻谷剛彦・志水宏吉，2004，『学力の社会学—調査が示す学力の変化と学習の課題』岩波書店。
- 荻谷剛彦，2008，『学力と階層』朝日新聞出版。
- 川口俊明，2011，「日本の学力研究の現状と課題」『日本労働研究雑誌』53(9), 6-15.
- 木村拓也，2006，「戦後日本において「テストの専門家」とは一体誰であったのか？：戦後日本における学力調査一覧と「大規模学力テスト」の関係者一覧」『教育情報学研究』4, 67-99.
- 国立教育政策研究所編，2010，『生きるための知識と技能—OECD生徒の学習到達度調査（PISA）2009年調査国際結果報告書』明石書店。
- Martin, M.O. & Mullis, I.V.S., 2012, *Methods and procedures in TIMSS and PIRLS 2011*, TIMSS & PIRLS International Study Center.
- 耳塚寛明，2007，「小学校学力格差に挑む」『教育社会学研究』第80集，23-39.
- 村山航，2006，「PISAをいかに読み解くか—求められる評価リテラシー—」東京大学大学院教育学研究科基礎学力研究開発センター編『日本の教育と基礎学力—危機の構図と改革への展望』明石書店，70-91.
- OECD, 2012, *PISA 2009 Technical Report*, OECD.
- OECD, 2010, *PISA 2009 Results*, OECD.
- 静哲人，2007，『基礎から深く理解するラッシュモデリング—項目応答理論とは似て非なる測定のパラダイム』関西大学出版部。
- 須藤康介，2013，『学校のエデュケーション効果と階層—中学生の理数系学力の計量分析—』東洋館出版社。
- 豊田秀樹，2002，『項目反応理論【入門編】』朝倉書店。
- 恒吉諒子，2008，『子どもたちの三つの「危機」』勁草書房。
- Wu, M., 2009, *Issues in Large-scale Assessments*, Keynote address presented at PROMS 2009, July 28-30, 2009, Hong Kong.
- 全国大学国語教育学会編，2010，『国語学力調査の意義と問題』明治図書出版。

