

京都府小学生の国語の学力を年度間比較するための調査

松 宮 功

本研究では、京都府教育委員会が作成して実施した小学校6年生対象の国語の学力テストによって測定される学力を年度間比較した。具体的には、1992・2001・2010（平成4・13・22）年度に行われた京都府の学力テストの既出項目から3セット6種類の調査テストを作成し、年度末を迎えた5年生709人を対象に調査を行った。調査で得られたデータと比較対象とした3か年の学力テスト受検者約37,360人のデータを使い、項目反応理論による等化を行った。その結果、比較した3か年の間で、小学校6年生4月時点の国語の学力特性値には変化が見られないことが示された。

キーワード：小学校6年，国語，学力テスト，経年変化，項目反応理論，等化，学習指導要領

1. 問題と目的

1.1. 国語学力の把握

PISA2000読解力リテラシーの国別順位低下が公表されて以後、国内にあった学力低下の論点は、PISA型読解力、国語力、言語力など、国語の教科学力を基盤とする学力に注目が集まるようになった。それは、2008年3月告示、2011年度完全実施の小学校学習指導要領改訂に少なからず影響した。改訂には、全教科において言語活動を積極的に取り入れることが盛り込まれ、それが重点の一つとなった。

国語の学力を把握する方法は様々あるが、学力問題の論議の際、その資料として何らかの学力テスト結果が頻繁に利用される。学力のすべてを表現することはできないものの、学力テストによる学力把握の方法も有効な一つである。

1.2. 学力の経年変化の把握における現状

テストは、多岐にわたる教科の内容の習得状況を把握するために、複数のテスト項目を用意して、受検者の反応情報を収集する。収集された正誤情報に得点が与えられ、それを合計することによって学力を測定する。すなわち数値で表現する。数

値になると比較しやすいから使いやすい。

しかし、異なる複数のテスト結果を比較したい場合、重大な問題が生じる。例えば、A年度のテストAを受けた乙の得点が70点で、B年度のテストBを受けた甲が75点だったとしても、甲の学力が乙より高いかどうかはわからない。テストAがテストBより難しいかもしれないからである。テスト項目の質が高かろうが、受検者が多かろうが、テスト結果から表現する学力を合計点によって運用している限り、この問題は解決しない。テスト難易度と受検者学力を分離できないからである。

この方法によってテスト結果から得た情報を学力として数値表現する場合、義務教育学校の児童生徒の学力を、経年比較可能な形で追いつけることは難しい。都道府県レベルの学力テストはもとより、全国学力・学習状況調査も同じ問題を抱えている。PISA（OECD）の学力の経年推移がしばしば示されるが、PISA型読解力が重要視されてきたという理由だけではない。PISA調査が、経年比較を可能にする運用がされているからである。学習指導要領を基準とした教育課程の下で学習する、国内児童生徒の学力を示すものとして、特定の学校カリキュラムがどれだけ習得されているかをみるものではないPISA（国立教育政策研究所、2002、

2004b, 2006b) の結果に頼らざるを得ないところに、日本における学力把握の課題が象徴されている。

一方、小中学校を対象とした都道府県レベルの学力テストは2002年以後急増し(文部科学省調査)、2007年には全国学力・学習状況調査が行われるに至った。学力テストの実施教科として、国語は常に一番手である。国・都道府県が効果的な教育施策を検証するには、学力の経年比較は欠かせないが、これらの学力テストは、経年比較を可能とする設計はされていない。

「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」(全国的な学力調査の在り方等の検討に関する専門家会議, 2011)に、義務教育学校の学力を経年で追う国レベルの仕組みの構築の必要性が記述された。続いて2012年1月、平成25年度以後のきめ細かな調査の実施が提起された。それを受け、テスト項目を一部非公開とした上で経年比較することを提案している(全国的な学力調査の在り方等の検討に関する専門家会議, 2012)。ただ課題は多く、具体的な制度設計はこれからである。

単独のテストデータが何万件あろうが尺度を共通化することはできない。尺度を共通化することを等化という。

PISAのように年度を越えて同一項目複数回提示し、異なるテストに使われたテスト項目固有の指標(難易度等)を同一尺度上に乗せる手法がある。これは共通項目計画と呼ばれるテスト運用である。大規模テスト運用の世界標準となっているこの手法が、これまで国内学力テストで実現していない理由は、テスト項目の初出性尊重主義(中畝・内田 2002)や項目使い捨て主義(豊田 2001)が日本社会の常識となっているからであるとされる。

1.3. 国内の先行研究

テスト項目を共通とする共通項目計画とは別に、異なるテストを比較可能とするため、テスト実施後に受検者を確保し、比較したいテストを受検してもらい、そのデータをもとに尺度を揃える共通受検者計画という方法もある。大規模学力テストのテスト項目公開が前提となっている我が国においては、この方法が現実的である。

この方法を用いた研究としては、前川・石塚ら(2001)、前川・菊地ら(2003)、吉村・荘島ら

(2006)、熊谷(2007)、斉田(2003)などがある。いずれも大学入学時または高校生を対象としている。この方法を用いて、小・中学生の学力の経年比較を試みた報告としては、宮崎県中学校教育研究会が行っている県数学一斉テストを用いた矢野・藤井(2008)、小学校6学年算数を対象とした国立教育政策研究所の調査を活用した長崎・荻原(2004)がある。しかしこれらは算数・数学であり、義務教育学校段階の国語の学力を潜在特性として経年比較した例は見当たらない。

国の将来を担う人材を育てる基盤が義務教育であり、すべての教科学力の基盤が国語であることは論を待たない。小中学校における国語科の学力を経年比較可能な形で把握することは極めて重要であると言えよう。

1.4. 本研究の目的

共通受検者計画は、受検者を共通尺度として過去のテストの難易度を推定するため、共通受検者の偏りによる等化への影響が問題になる。従って、教育課程の基準たる学習指導要領の下で指導が展開される小学校においては、改訂の影響は不可避である。この点が、経年比較をさらに難しくしている理由の一つであろう。しかし、学習指導要領改訂を含む長期間にわたる小学生の学力の年度間比較は、義務教育であるがゆえに、教育課程や教育施策の検証という観点から重要であり、それを試みることは意義深い。

本調査は、1992・2001・2010年度の小学校6年生を対象とした京都府小学校基礎学力診断テスト(以後、京都府学力テスト)の国語の結果を共通受検者計画によって等化し、小学校6年生4月時点での国語を年度間比較し、学習指導要領改訂による授業時間数の変遷や教育施策の観点から考察することを目的とする。

2. 京都府学力テストと比較対象年度

2.1. 比較対象年度

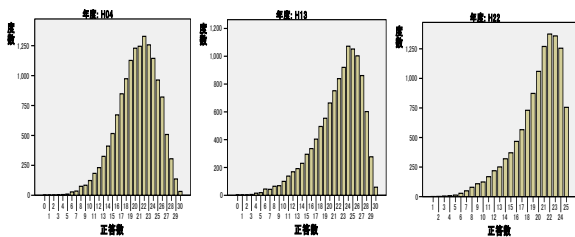
京都府の学力テストは、府内公立小学校(京都市を除く)の4年生・6年生が対象である。前学年までの学習内容を出題範囲として4月に実施される。児童全員の受検を前提とするこのテストは、学校設置者の希望制ではあるが、過去参加しなかった小学校はなく、事実上悉皆である。

1992（平成4）年を比較初年度として、9年間隔で2001・2010（平成13・22）年度とした。これによって対象年度の受検者は、入学時からそれぞれ、小学校1年生から5年生までの5年間を、昭和52・平成元・平成10年告示年度板学習指導要領の下で学んだことになる。

分析対象とした学力テストの受検者数と項目数を表1に示す。図1は年度別・教科別の度数分布である。

表1 比較対象年度の学力テスト

年度	1992 (H4)	2001 (H13)	2010 (H22)
項目数	30	30	25
受検者数	14,622	11,281	11,457



<1992年度> <2001年度> <2010年度>

図1 正答数の度数分布

2.2. 項目反応理論と等化

本研究の共通受検者計画は、新たに確保した共通受検者の正誤パタンのデータを基に、欠測データを含めた分析対象全体の項目反応状況を利用し、受検者の学力値と項目固有の特性（難易度など）を推定する。そのとき適用するテスト理論が項目反応理論（IRT；Item Response Theory，豊田 2002 など）である。この理論のモデルは複数あるが、本研究では3パラメタ・ロジスティックモデル（3PLM）を用いた。3PLMは、学力特性値が θ である受検者が項目に正答する確率を

$$P(\theta|a,b,c) = c + \frac{1-c}{1 + \exp\{-a(\theta-b)\}}$$

とする（荘島 2010）。但し、 a ：傾き母数、 b ：位置母数、 c ：下方漸近母数である。 a は識別力、 b は困難度、 c は疑似チャンスレベルとも呼ばれ、項目固有の値である。

図2は項目母数によって表現される項目反応関

数の例である。3PLMでは、 a は曲線の傾きをつかさどり、 b は曲線の位置を決める。 b は正答確率0.5となる受検者の学力特性値 θ である。問題難易度は受検者学力の尺度で表現される。難しい問題はグラフが右に、易しい問題は左に寄る。 c は左すその高さを表す。例えば、記述や作図問題では、 c は0に近い値となる。グラフが著しく左に位置する非常に易しい問題の場合、 c は高めになる。

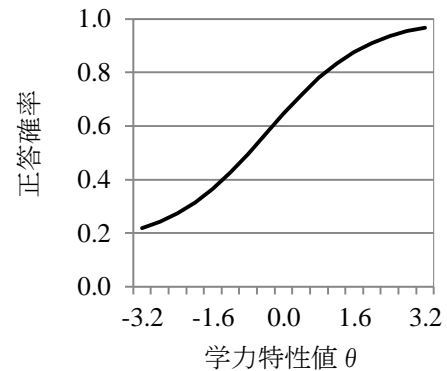


図2 項目反応関数（平成4年度No.2）

また a は、 b 値付近の受検者学力帯の識別力の程度を表す。 a が高いと、傾きが急となる付近の学力帯を識別する情報量を多く持っていることを意味する。難しい問題は学力高位層を識別する情報を持つが、低位層を識別する情報はほとんど持たない。すなわち、項目のテスト情報量は学力帯によって異なる関数になる。項目全ての情報量関数を積算した関数が、テスト情報関数（図3）と呼ばれ、テスト全体の学力帯による学力測定力を表している。

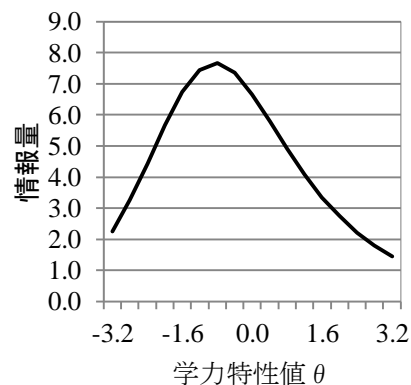


図3 テスト情報関数（等化後）

2.3. 京都府学力テストと学力

京都府学力テストの目的は、「学習指導要領に示す目標や内容に照らした学習の実現状況を（中略）、一人一人の生徒に基礎・基本の確実な定着による学力の充実・向上を図る」（平成22年度）である。学習指導要領の示す内容の定着状況を把握するテストと考えてよい。

テスト時間は45分、解答形式は4肢選択問題が多く、マークシートを採用している。誤答は陥りやすいパターンによって作成されている。記述形式の項目は、詳細な採点基準に従い、指導者によって採点され、誤答類型別にマークされる。

多肢選択式が記述式かによって問題の優劣は論じられない（日本テスト学会，2007）。解答形式の違いによって一長一短があるというほうが適切である。

記述形式が増えると、テスト時間に制限ある限り問題数を減らさざるを得なくなる。問題数が減れば、広い範囲を代表するテストとして妥当性が低くなる。多肢選択式を採用して項目数を増やすことによって、テスト全体の妥当性や信頼性は高まる。また、記述形式による解答を求めた場合、その採点者の採点基準に対する解釈のゆらぎが誤差となり、客観テストとしての質を保つことを難しくする。

しかし一方では、選択式では当て推量によって正答することが問題になる。この点、京都府学力テストは丁寧に検討されている。蓄積された過去問題の反応状況を参考に、児童が陥りやすい誤答パターンを考慮して、誤答選択肢が用意される。記述問題にあっても、提示された誤答類型に従って指導者がマークすることで区別する。当て推量による正答は、IRT分析においてはc(下方漸近母数)値を指標としてある程度評価される。従って、当て推量の要素を含めた上で、項目特性や受検者学力が推定されている。

一般にテストが大規模になればなるほど、実施可能性から多肢選択式を増やさざるを得ない。記述形式にすると、豊かな解答情報を収集できるが、種類の多いデータを収集できず、テスト結果の一般性が得られにくくなる。それはトレードオフの関係にあると言ってよい。

記述形式によって得られる情報は、教師作成テ

ストやノートを通じて、児童（受検者）を直接指導する教員が日常的に得ることが望ましい。

3. 調査

3.1. 調査項目の選択と調査テスト冊子の作成

2010（平成22）年度学力テストにはリスニングが3項目含まれている。調査テストの項目選択に当たっては、まずリスニング問題を除外し、採点者のバイアスの混入が少ない選択形式に絞った。また、実施当時と同じ状況を再現するために、大問はまとめて選択した。最終的に、各年度とも13項目が選択された（表2）。

なお、調査項目は、調査テスト冊子内に学力テストの実施年度と同じ順序で配置された。

選択された項目群セットは、3か年分3種類（以後、1992年度：セットA、2001年度：セットB、2010年度：セットC）ある。この内、2セットを組み合わせて調査テスト冊子を、AB、BA、AC、CA、BC、CBの6種類作成した。AB・BA、AC・CA、BC・CBの逆転配置は、順序効果を相殺するためである。

表2 調査テストに採用した項目数

年度:セット	調査項目数	実施年度テスト (リスニング)
1992 : A	13	30(-)
2001 : B	13	30(-)
2010 : C	13	25(3)

3.2. 調査協力者と実施

京都府内（京都市を除く）公立小学校約200校から地域や規模の均衡に配慮して16の調査協力校が選ばれた。

調査は、2010年度の小学校5年生を対象として、2010年1～3月に行われた。調査実施期日は学校に任された。調査時間は、オリジナルテストと同じ45分で実施された。共通受検者数を表3に示す。

なお、1つの調査協力校に6種類の内1種類だけを配布すると、その学校規模や地域によるデータの偏りを生じる可能性がある。これを回避するために、6種類の調査テスト冊子すべてが、調査協力校のすべての学級において、無作為に配布された。

表3 調査参加者数

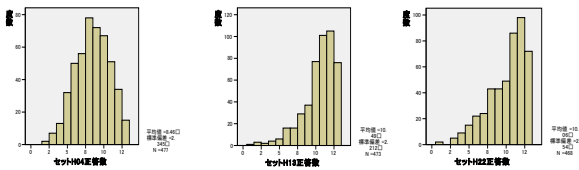
冊子	AB	BA	AC	CA	BC	CB	計
人数	120	121	119	117	119	113	709
セット別							1992年度：セットA 477
							2001年度：セットB 473
							2010年度：セットC 468

3.3. 調査結果の概要

調査テスト正答数のセット別要約統計量とセット間相関係数を表4に、セット別の正答数の度数分布を図4に示す。

表4 調査テスト正答数の要約統計量

年度:セット	正答数				相関係数		
	Mean	SD	Max	Min	A	B	C
1992 : A	8.46	2.35	13	2	1		
2001 : B	10.49	2.21	13	1	.56	1	
2010 : C	10.06	2.54	13	1	.57	.74	1



1992:セットA 2001:セットB 2010:セットC

図4 調査セット別の正答数の度数分布

4. 結果

国語は、7項目を除く78項目を分析対象とした。分析から除外した項目は、2010（平成22）年度のリスニング3項目、識別力が著しく低かった4項目（1992年度2項目、2001年度1項目、2010年度1項目）の計7項目である。

表5 項目母数の要約統計量と標準誤差の平均
国語（対象78項目、1991:28, 2001:29, 2010:21）

項目母数	Mean	SD	Max	Min	平均SE
a : 傾き	0.834	0.262	1.462	0.337	0.024
b : 位置	-0.888	1.214	2.064	-4.187	0.067
c : 下方漸近	0.257	0.124	0.645	0.008	0.017

推定された特性値 θ （学力）は、比較初年度である1992（平成4）年度受検者の平均値が50、標

準偏差が10となるように換算した尺度得点を使って比較した（表6、図5・図6）。図5の平均値の上下に延びる直線は標準偏差である。

表6 特性値 θ と尺度得点

年度	θ Mean	尺度得点		
		Mean	SD	平均SE
1992 (H4)	0.172	50.00	10.00	5.00
2001 (H13)	0.298	50.86	10.13	4.57
2010 (H22)	0.334	51.10	10.10	5.23

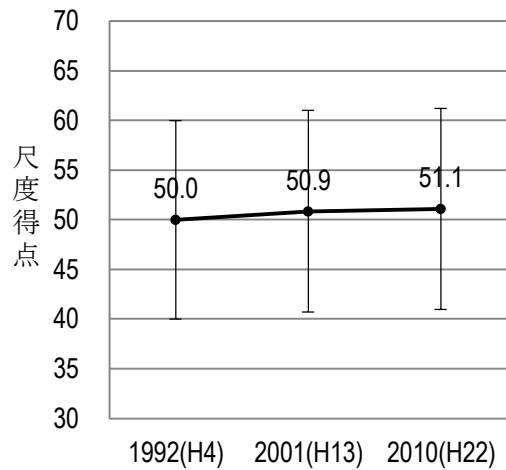


図5 国語学力の経年推移

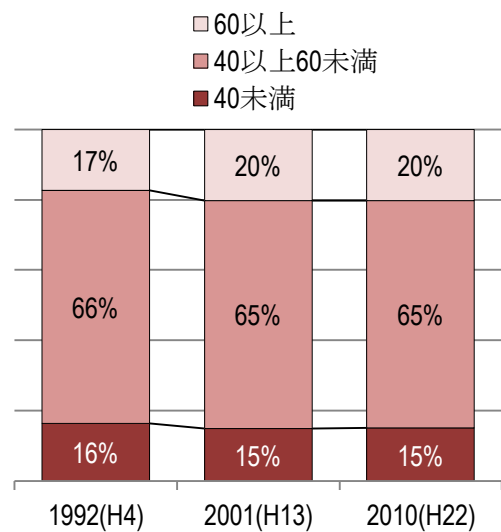


図6 尺度得点 3層の経年推移

5. 考察

尺度得点平均値は、1992年度から18年経過した2010年度においても大きな変化は見られない（図4）。

この結果をもって、学力は18年前と比較して低下していない、という一般化された結論を出すことは危険である。尺度得点の表す学力は、あくまでも京都府の学力テストの測定する学力に限定されるからである。ただ、京都府の学力テストが学習指導要領を基にして基本的な内容を出題していることを考え合わせると、京都府の6年生に限れば、少なくとも18年間基礎的な国語の学力は低下していないと考えられる。

児童の学力テスト受検まで（小1～小5）の国語科の授業時数を整理すると表7のようになる。

表7 比較対象年度の小学校6年生が
小1～小5に受けた国語科の授業時数

生誕年度	入学年度	週5日制	小1～小5 授業時数	受検年度
1980	1987	週6日制	1,322	1992
1989	1996	隔週5日	1,391	2001
1998	2005	週5日制	1,202	2010

尺度得点が18年前に比べて下がっていないという結果は、予想を覆すものだった。なぜなら、この間の学校週5日制への移行と改訂による授業時数の減少は、学力高位層は別にして、中低位層への負の影響は避けられないと推測されたからである。では、なぜそうならなかったのか。2001年度と2010年度に分けて、学習指導要領改訂による授業時間数の変遷や教育施策の観点から考察する。

まず、2001（平成13）年度である。1992（平成4）年度との間では、学校週6日制と隔週5日制の違いはあるが、学習指導要領改訂による授業時数の減少はない（表7）。この環境下で尺度得点がやや上昇した理由として、学力テストの継続実施による効果が考えられる。この9年間で、京都府全域の小学校で、自校のテスト結果を踏まえた校内研修が実施されるに至った。これは指導者が、理念的・抽象的な表現ではなく、テスト問題という形で、日々の授業の具体的な達成目標を提示され、それを達成するためにどのような指導を行えばよいかを振り返る機会になったと言えるだろう。言い換

えれば、テスト問題が、学校教育現場に対して、指導の方向性を示したことに他ならない。全国学力・学習状況調査の問題が、学校教育現場に強いメッセージ性を持った効果と共通すると考えられる。

それを後押しすることとして、京都府総合教育センターが毎年開催したテスト分析・活用講座がある。そこでは、詳細な資料（学力診断テストの概要、毎年度発刊）が配布され、分析結果を踏まえた校内研修会の持ち方が協議された。その内容は、「京都府小学校基礎学力診断テスト：10年間のまとめと学力充実の方策」（京都府教育委員会、2001）に詳しい。これらの影響によって、広域で一定の指導水準が保たれたことが、尺度得点の差となって表れた理由の一つと考えられる。

次に2010（平成22）年度である。この年度の6年生は、完全学校週5日制下の児童であり、減少した授業時数で小学校5年間を過ごしている（表7）。それでも、比較初年度である1992（平成4）年度6年生と比べて尺度得点は高く、2001（平成13）年度6年生と比べても、同程度の水準を維持している。この結果に明快な理由を特定することは難しい。それでも、教育施策の観点から少なくとも2つの理由が考えられる。

第1は、指導方法の工夫改善を目的として行われた加配措置である。これは、公立学校の教職員数を規定する義務標準法において、学級数等に基づく基礎定数とは別に、加配定数として配置される枠で、指導方法の工夫改善を目的とした増員である。2001（平成13）年度から始まったこの措置は、単なる増員に留まらず、加配教員が学校全体の国語・算数の教科指導を牽引する活用が強調され、実際に学校はそのような活用を意図した。加配教員の授業計画・教材準備によって、校内の指導者全員の指導水準が上がったことが考えられる。

第2は、この人的配置の京都式の運用である。加配教員の具体的活用方法は、ティームティーチング（TT、1学級2教員の授業）、少人数授業（例えば、2学級3グループ展開）、少人数学級（1学級の人数を少なくするために、加配教員を担任として1学級増やす）などがある。京都式の運用（「子どものための京都式少人数教育」と呼ばれる）の特長は、設置者・学校がその活用方法を選択できることにある。この結果、地域の実情、学校規模、

児童生徒の実態，教員構成等を考えて，児童にとって最も効果的だと学校が判断する活用方法が選択され，効果的であったと考えられる。

学級編制と少人数指導形態が小学校4学年時と6学年時の学力に与える影響を明らかにすることを目的とした，国立教育政策研究所の研究報告書（工藤，2012）によれば，「少人数指導実施形態の違いでは説明されないが，国語において，従前の学力が同程度の児童でみると，現行の基準によって編制された30人を超える学級規模の学校の児童より，現行を下回る基準による学級編制を継続的に実施した児童の方が，その後の学力が高い」と，「現行を下回る基準による学級編制を継続的に実施した学校に在籍することが，従前の学力が低い児童に対して補償的である」ことが示された。この報告書の調査対象は，「京都府内（京都市を除く）の小学校のうち，2011（平成23）年度6学年において単式学級が2以上あった110校（8120人）」（同報告書）であり，本研究の等化用調査参加者が含まれる。また使用された学力データは，本研究が対象としている京都府学力テスト（2010年度4年，2012年6年）である。

この結果を合わせると，指導方法工夫のための加配教員の配置が効果をあげたことが示唆される。

6. 本研究の限界

学力テストが国語の学力の一部しか測定していないこと，調査が学習指導要領の改訂をまたぐ18年間に及ぶため，対象のテストが測定している学力の範囲が完全には一致していないこと，対象テストの年度が3か年であることから，尺度得点のわずかな上昇をもって，国語力が向上したと結論づけることはできない。正確には，「1992（平成4）年度6年生にできていたことは，2010（平成22）年度6年生も同程度以上できる」と表現するのが適切であろう。

さらに，テスト受検者の範囲が京都府内に限られていること，学習指導要領改訂や人的配置は全国共通であっても，加配教員の運用の仕方や国語力向上の施策には地域性があることから，今回の結果を，国内全域の小中学生に一般化することは慎まなければならない。

謝辞

本研究の調査にご協力いただいた京都府内の市町教育委員会と小学校に感謝いたします。

また，本研究にあたり，調査設計の段階から分析に至るまで助言いただいた荘島宏二郎先生（独立行政法人・大学入試センター）に感謝いたします。

参考・引用文献

- 荒井克弘・倉元直樹編著（2008）全国学力調査日米比較研究，金子書房
- 池田央（1992）テストの科学—試験にかかわるすべての人に—，日本文化科学社
- 池田央（1994）現代テスト理論，朝倉書店
- 熊谷龍一（2007）大規模英語学力テストにおける年度間・年度内比較—大学受験生の英語学力の推移，日本テスト学会誌，**3**，pp.83-90.
- 加藤幸次編著（2001）タイプ別学習集団の効果的な編成，ぎょうせい
- 京都府教育委員会（2001）京都府小学校基礎学力診断テスト：10年間のまとめと学力充実の方策，平成13年12月
- 工藤文三（2012）学級編制と少人数指導形態が児童の学力に与える影響についての報告書，国立教育政策研究所，p.3
- 国立教育政策研究所（2002）生きるための知識と技能—OECD 生徒の学習到達度（PISA）2000年調査国際結果報告書，ぎょうせい
- 国立教育政策研究所（2004）生きるための知識と技能2—OECD 生徒の学習到達度（PISA）2003年調査国際結果報告書，ぎょうせい
- 国立教育政策研究所（2007）生きるための知識と技能3—OECD 生徒の学習到達度（PISA）2006年調査国際結果報告書，ぎょうせい
- 前川眞一・石塚智一・菊地賢一・内田照久・中畝菜穂子（2001）大学入試センター試験得点の標準化の試み，大学入試研究ジャーナル，**11**，pp.15-23.
- 前川眞一・菊地賢一（2002）大学入試センター試験得点の標準化の試み—項目反応理論による方法—，平成11年～13年度共同研究報告書「大学入学者選抜における評価の標準化に関する

研究」(研究代表者 石塚智一), pp.33-72.

中畝菜穂子・内田照久(2002) 大学入学者選抜資料の評価の標準化に関する意識調査—高校生, 高校教員, 大学教員の意見—, (石塚智一研究代表者) 大学入学者選抜における評価の標準化に関する研究, 大学入試センター研究開発部, pp.19-31

長崎栄三・萩原康仁(2004) 算数達成度の項目反応理論による比較分析, 国立教育政策研究所

日本テスト学会編(2007) テスト・スタンダード—日本のテストの将来に向けて, 金子書房

日本テスト学会編(2010) 見直そう, テストを支える基本の技術と教育, 金子書房

斉田智里(2003) 高校入学時の英語能力値の年次推移—項目反応理論を用いた県規模英語学力テストの共通尺度化—, *Step Bulletine*, **15**, pp.12-24

荘島宏二郎(2010) 項目反応理論—学力を測定するためのテストの科学, 植野真臣・荘島宏二郎(編著) 学習評価の新潮流 朝倉書店, pp.56-82.

Shojima, K. (2010) *Exametrika 4.4* (<http://www.rd.dnc.ac.jp/~shojima/exmk/index.htm>)

高浦勝義(2004) 指導方法の工夫改善による教育効果に関する比較調査研究—授業法の違いが児童生徒の学力, 興味・関心・態度及び学習態度の形成に及ぼす効果について(第二次・最終報告書)—文部科学省科学研究補助金・研究成果報告書

植野真臣・永岡慶三(2008) e-テストィング, 培風館

矢野愛子・藤井良宜(2007) 宮崎県数学一斉テストにおける中学生の学力の変化についての分析, 第40回数学教育論文発表論文集, pp.61-66

山森光陽(2006) 学力低下論争, 目標標準評価の定着, 学力テストブームの狭間で, 教育心理学年報 2006年 第**45**巻, pp.92-103

吉村幸・荘島宏二郎・杉野直樹・野澤健・清水裕子・齊藤栄二・根岸雅史・岡部純子・サイモン フレーザー(2006) 大学入試センター試験既出問題を利用した共通受験者計画による英語学力の経年変化の調査, 日本テスト学会誌, 1-1, pp.51-58

全国的な学力調査の在り方等の検討に関する専門

家会議(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, 平成 23 年 3 月 31 日

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

全国的な学力調査の在り方等の検討に関する専門家会議(2012), 「きめ細かい調査」の基本的な枠組み」の公表について, 平成 24 年 1 月 27 日

(http://www.mext.go.jp/b_menu/shingi/chousa/s-hotou/085/houkoku/1316096.htm)