第 3 章
「JHPS お子様に関する特別調査 2010」
――学力テストの信頼性と妥当性の検討

敷島千鶴
直井道生
山下 絢
赤林英夫

要旨

「JHPS お子様に関する特別調査（子ども特別調査)」は、日本家計パネル調査(JHPS) 対象者の中から小中学生の子どもをもつ親を抽出し、その親子を JHPS に付随して追跡することにより、子どもの学力と家庭背景との関連を解明することを目的とした調査である。本稿では、2010 年に実施された第 1 回目の調査において収集された、312 世帯の子ども 461 名分のデータを分析することにより、子ども特別調査学力テストの信頼性と妥当性について検討を行った。算数/数学、国語、推論の 3 因子に従属する項目群の内的一貫性の高さは、子ども特別調査学力テストが優れた信頼性を伴うことを示した。都道府県別学力テストデータとの比較は、子ども特別調査の学力テスト各項目が構成概念妥当性を伴うことを示した。算数/数学、国語のテスト得点と親が報告した子どもの学校の両教科の成績との相関関係は、子ども特別調査学力テストで測定される構成概念が併存的妥当性を伴うことを示した。子ども特別調査学力テスト得点が、有効な学力指標であることが確認された。

第１節　はじめに[1]

子どもの学力と家庭の背景との関係に関心が高まってきている。社会学、経済学における議論の趨勢は、子どもの学力は家庭の社会経済的背景によって規定されると主張し（苅谷・志水, 2004；志水, 2005；耳塚, 2007；橘木・松浦, 2009 など)、現代日本の世論に対し、教育格差とその再生産という問題意識を提起している。

しかしながら、我が国において、こうした議論の多くがこれまで依拠してきたのは、学校を通したクロスセクショナルな調査結果であるか、もしくは成人のみを対象とした社会調査の分析結果である。つまり、子どもの学力に関する膨大なデータと、家庭背景に関する膨大な情報群は、うまくリンクされていない。全国を母集団とする偏りの無いサンプルを対象に同一世帯を追跡し、子どもの学力と家庭の社会経済的情報の双方を詳細に収集することにより、両者の関係性に対する実証的知見を提供していくことは、喫緊の課題である。

こうした背景のもと、成人を対象とした家計パネル調査の中に、その家庭の子どもの学力を実際に測定するテストを組み込むという試みがなされている。それが、「JHPS お子様に関する特別調査」(以下、子ども特別調査という）である。このデザインを用いれば、両者で測定される変数間の静的な記述にとどまらず、家庭背景と子どもの学力の間の動的な相互関連性を解明することが可能となる。

子ども特別調査の目的は、同じ子どもの学力と生活状況を追跡することにより、家庭の社会経済状態と子どもの教育達成度の因果関係を解きほぐすこと、National Longitudinal Survey of Youth（NLSY）に近い設計をすることにより、米国における研究と比較可能にすること、そして、日本家計パネル調査（JHPS）の豊富な家計情報を利用することにより、これまで気づかれなかった、家庭背景と子どもの教育達成度との関係を明らかにする材料を提供することにある（JHPS 子ども特別調査ワーキンググループ, 2010: p7)。

本稿は、こうした目的を遂行すべく、子ども特別調査を展開していく前段階として、その基軸となる子どもの学力測定のために独自に開発した学力テストに焦点をあて、その信頼性と妥当性について検討する。テストの信頼性とは、そのテストが、測定しているものをどの程度安定し、一貫して測定しているかについての評価である[2]。一方、テストの妥当

---

[1] 本稿の一部は、2010 年 11 月、慶應義塾大学にて開催されたパネル調査共同研究拠点ワークショップにおいて報告された。ワークショップでの報告に際し、貴重なコメントを下さった河合啓希先生（慶應義塾大学経済学部)、「JHPS お子様に関する特別調査」実施にあたり、ご尽力いただいた中村亮介氏（慶應義塾大学大学院経済学研究科)、上村一樹氏（慶應義塾大学大学院経済学研究科)、相澤佐知氏（慶應義塾大学経済学部)、飯崎　尭氏（慶應義塾大学大学院経済学研究科)、荒木宏子氏（慶應義塾大学大学院経済学研究科)、推論問題開発を手掛けてくれた佐藤有理氏（慶應義塾大学大学院文学研究科)、構造方程式モデリングを用いたデータ解析に助言を与えてくれた山形伸二氏（慶應義塾大学先導研究センター）に深謝申し上げる。

[2] 信頼性を表す指標として信頼性係数が用いられることが多いが、信頼性係数は、テスト得点の分散 σ^2_X を、真値による分散 σ^2_T と、誤差変動による分散 σ^2_E の和で表現した時に、比 σ^2_T/σ^2_X で定義される。推定

性とは、そのテストが、目的とする測定すべき構成概念をどの程度正しく反映しているかについての評価である[3]。測定の道具である学力テストを、分析の１つの変数である学力指標として機能させるためには、そのテストが、精度高く、測定したい対象を正確に捉えていること、すなわち、そのテストが高い信頼性と妥当性を備えている必要があることはいうまでもない。本稿ではこのことを確認するために、以下の３つの検討を行った。

第一は、テスト項目の信頼性の検討であるが、ここでは、まず、それぞれのテスト項目群がどのような相互関連性をもつのか、その構造を確認する。子ども特別調査学力テストは、小学１年から中学３年まで各学年、算数/数学、国語、推論の設問で構成される。算数/数学、国語の項目群については、それぞれの科目の学習指導要領の中で作成された問題から選択されていることから、内容的妥当性を十分満たしていると考えてよい。また、推論の項目群についても、論理学研究者が推論問題として開発したことより、その内容的妥当性はクリアしていると考えられる。しかし、それぞれの科目が測定する構成概念間の構造については、検討の余地がある。算数/数学、国語、推論のそれぞれの項目群が、相関はあっても､異なる構成概念を測定しているということを示すためには､それぞれの項目群が、背後にある算数/数学、国語、推論という３つの潜在因子で表現される構成概念から、それぞれ負荷を受けることを仮定する３因子モデルの当てはまりがよいことを示す必要がある。本稿では、モデルの当てはまりを評価できる構造方程式モデリングのモデル適合の方法を用いて、確証的因子分析を行うことにより、テスト項目の因子構造を確認する。

そして、因子構造に基づいた項目群の信頼性を、テストの内的一貫性の立場から評価する。内的一貫性とは、同一個人が同様のテスト問題に対し、同じように回答している程度であり、α 係数という反応の整合性あるいは等質性を数値化した指標を用いることによって検討できる。

第二は、既存データとの正答率の比較に基づく、学力テスト項目レベルでの構成概念妥当性の検討である。具体的には、子ども特別調査学力テストと、子ども特別調査に先がけて実施した予備調査（JHPS 子ども特別調査ワーキンググループ, 2010)、ならびに小学 4 年から中学 3 年までの算数/数学､国語の問題の多くの出典となった都道府県別学力テストの間で、それぞれ対応のある項目の正答率を 2 種類の方法で比較する。1 つめの方法は、平均値の差の検定であり､サンプルによるパフォーマンス水準の差についての検討である。都道府県別レベルで実施された大規模調査と子ども特別調査のサンプルの学力が、共に全国の水準を代表したものであるならば、両サンプルに対して施行されたテスト項目の正答率には差が無いはずである。そしてもう一つの方法は、２つの調査間で測定された複数の

方法には、再テスト法、折半法、平行テスト法、内的一貫性による方法などがある。

[3] 妥当性は、内容的妥当性、基準関連妥当性、構成概念妥当性に大別される。内容的妥当性とは、テストの内容が関心とする測定領域をどれほどよく代表しているかを、テスト作成者や専門家が直観的、経験的に判断するものである。基準関連妥当性は、客観的な外的基準との相関関係によって評価されるが、同時に測定された、測定対象が同じである他の変数との相関関係をとくに併存的妥当性と呼ぶ。構成概念妥当性は、研究や理論構成に必要な概念、すなわち構成概念をどの程度測定しているかを評価する包括的な概念である。

項目群の正答率についての相関係数の算出であり、項目間の共変動の一致についての検討である。相関係数が 1 に近ければ、２つの調査の項目間の共変動の一致の程度は高いことが示され、テスト項目の構成概念妥当性は、やはり高いと判断できる。

そして第三は、学力テストで測定された構成概念の妥当性の検討である。子ども特別調査学力テストの科目別得点と、その個人の学校の科目別成績との相関関係を求めることにより、併存的妥当性が検討できるが、算数/数学の学力テスト得点と算数/数学の学校の成績が､そして国語の学力テスト得点と国語の学校の成績が､それぞれ高く相関していれば、収束的妥当性があると考えることができる[4]。また、それらの相関が、算数/数学の学力テスト得点と国語の成績、あるいは国語の学力テスト得点と算数/数学の成績との相関より高ければ、弁別的妥当性があるとも考えられる。

本稿では、これらの３つの視角から、子ども特別調査学力テストの信頼性と妥当性について検討していく。

第２節　方法

１　協力者

日本家計パネル調査（JHPS）は、全国に居住する 20 歳以上の成人男女が構成する母集団より､層化 2 段無作為抽出法により抽出された標本約 4000 名とその配偶者を対象とし、同世帯をフォローアップすることにより、我が国の経済行動の動的変化の解明を目指したパネル調査である（樋口他, 2010)。調査は 2009 年以来、毎年継続して行われ、第 2 回目の調査（JHPS2010）が 2010 年 1 月から 3 月にかけて実施された。

子ども特別調査は、この JHPS2010 の一環として、JHPS2010 調査票において、義務教育段階である小学校あるいは中学校に就学する子どもがいる JHPS 対象者に対し調査協力を呼びかけ、それに応じた親子を対象としている。調査は小学 1 年から中学 3 年の子ども個人を単位とし、調査票二部から構成された。一部は、子どもが回答する、子どもの学年に対応した学力テスト、及び学びについてのアンケート調査である。もう一部は、JHPS 対象者である親が回答する、子育てや子どもの学びについてのアンケート調査であり、親は子どもひとりにつき一部を回答する。

JHPS2010 に協力した対象者のうち、子ども特別調査対象年齢に該当する、小学校あるいは中学校に就学する子どもがひとり以上いるケースは 644 であった。そのうち、子ども特別調査への有効な協力数は 312 であった。ひとりの JHPS 対象者から協力した子どもの人数は、1 名が 180 ケース、2 名が 116 ケース、3 名が 15 ケース、4 名が 1 ケースであった。JHPS2010 に協力した対象者の子どもで、子ども特別調査対象年齢に該当する、小学校あるいは中学校に就学する子どもの人数は 959 であった。そのうち、子ども特別調査へ

[4] 同じ構成概念を測定している別の心理検査と正の相関関係を示すことを収束的妥当性、理論的に相関関係の低い別の構成概念と正の相関関係を示さないことを弁別的妥当性といい、共に構成概念妥当性の評価方法のひとつと考えられている。

表３－１　子ども特別調査学力テスト有効回答数

小1	小2	小3	小4	小5	小6	中1	中2	中3	合計
62	44	63	47	62	38	57	46	42	461

の有効な協力数は 461 であり、学年ごとの内訳は表３－１にある[5]。

子ども特別調査は、JHPS2010 調査票回収後、調査協力に同意した世帯に調査員が出向き、子どもと親の調査票と謝礼（図書カード 500 円分とクリアファイル）を渡した後、各家庭において実施された。

子どもの調査票にある学力テストは、親、あるいは高学年では子ども本人が 20 分間計測し、子どもが時間内にひとりで回答すること、そして、子どもが回答した調査票は、速やかに子ども自身が同封されたシールで４箇所を封緘した後、親に渡すことを協力者に求め、対象者への調査依頼文書と、子どもの調査票のフェースシートにそのことを明記した。子どもの調査票と親の調査票は、返信用封筒に入れて郵送回収された。

２　テスト項目

子ども特別調査がモデルとするのは、アメリカの NLSY において、対象者の子どもに対し２年に一度実施している NLSY-Child Assessment である。子ども特別調査学力テストの作成についても、この NLSY-Child Assessment を参考にした。

学力テストの科目は、小学 1 年から中学 3 年まで一律に、算数/数学、国語、推論の 3 科目とした。このことは、NLSY-Child Assessment において 1986 年以来、5～14 歳児を対象に、Peabody Individual Achievement Test (PIAT) から Math、Reading Recognition、Comprehension の学力テストを、WISC-R から短期記憶（Digit Span）を測定する知能テストを施行していることに倣っている。加えて、子ども特別調査では、NLSY のように面接法ではなく留置き郵送法を用いるという制約より、回答者にあまり負担を課さない時間内で、家庭におけるセルフ・アドミニストレーションに頼らざるを得ないという限界、及び採点のしやすさを考慮した上で、実現可能性を検討した結果でもある（詳細は JHPS 子ども特別調査ワーキンググループ, 2010 を参照のこと)。

算数/数学の問題は、計算と、数や図形操作に関する文章題から構成した。国語の問題は、語彙と漢字の読み書きから構成した。知能テストに関しては、論理的推論に関する基礎的な問題から構成した。どの学年も合計 20 分以内で回答できる分量にすることとした。

算数/数学、国語に関しては、学年ごとに学習指導要領に即した異なる問題を用意した。小学１年から小学３年までは、学習ドリル製作会社に希望を伝え、オリジナル問題の作成を依頼した。小学４年から中学３年までは、都道府県で独自に実施された都道府県別学力テストから、問題を抽出して利用した。都道府県別学力テストの問題とそれぞれの正答率

[5] 子ども特別調査の対象者であることが確認できず、抽出エラー（本来は JHPS 対象者の子どもが回答するべきところを弟が回答したなど）によって回答した可能性のある子ども４名を含む。

表３－２　予備調査有効回答数

	小1	小2	小3	小4	小5	小6	中1	中2	中3	合計
人数	50	64	67	53	34	61	26	32	32	419
学級数	2	2	2	2	1	2	1	1	1	14

については、自治体に情報公開を求めることによって閲覧することができる。秋田県で平成 17 年～20 年小学 4 年～中学 3 年に、新潟県で平成 16 年と 18 年小学 4 年～中学 2 年に、岐阜県で平成 15 年～20 年小学 5 年～中学 2 年に、香川県で平成 18 年～21 年小学 3 年～中学 3 年に実施したテスト問題を入手することができたため、各学年の算数/数学、国語の問題プールを作り、その中から適切と思われる問題を選定した。

推論問題に関しては、論理学研究者の協力を得てオリジナル推論問題を開発し、これを知能測定の代替指標とすることとした。知能の定義には諸説あるが、推論能力がその中核にあることは、知能研究者の間で合意が得られている（Gottfredson, 1997 ; Snyderman and Rothman, 1988)。とりわけ、前提となる 2 つの命題から、結論となる 1 つの命題を導く推論形式である三段論法の解法に反映される演繹的論理推論能力が、一般知能と深く関連すること（Shikishima et al, 2009)、自記式質問紙に組み込んだ三段論法課題数問が、知能を予測し得ることが報告されている（Shikishima et al, in press)。9 歳になれば、三段論法解法能力の測定が有効であることが知られているため（Bara et al, 1995)、子ども特別調査では、小学 4 年以上については、三段論法形式の論理推論問題を導入することにした。そして、小学 4 年から中学 3 年まで、正解を選択肢の中から選ぶ形式の共通の 4 問を用意したが、小学 4 年～小学 6 年については、各設問に解法を補助するための図形表象を伴わせた。小学 1 年～小学 3 年については、三段論法形式ではないが、3 つ名辞の関係性を表す 2 つの命題から、それらの相互の関係性を演繹的に論理推論し、正解を選択肢の中から選ぶ関係推論の 4 問を設けた。

作成された小学 1 年から中学 3 年までのテスト問題は、静岡県内の小学校 1 校と、中学校 1 校の協力を得て、2009 年 12 月、予備調査として実施された。予備調査の有効回答数は 419 であり、学年ごとの数の内訳は表３－２に示した。

予備調査の目的は、作成したテストを実際の学校の教室で実施することによって、問題の難易度、設問形式の適切性、分量ならびに所要時間の適切性を確認することにあった (JHPS 子ども特別調査ワーキンググループ, 2010)。各学年の収集されたデータは、まず、項目ごとに正答率を算出し、正答率が著しく高く、分散が極めて小さい項目を削除した。科目ごとに正解を 1 点、不正解を 0 点とした単純加算得点の分布を調べ、天井効果のある科目については、問題に修正を加える、あるいは、都道府県別学力テストで公表された正答率がより低い新たな問題を挿入するなどの措置を施すことにより、難易度の上昇を試みた。さらに、各学年、科目ごとに内的一貫性を低める項目を削除した。全員から入手した学力テスト回答に要した時間の情報をもとに、各学年の問題量を回答時間 20 分間に相当すると判断できる分量に調整した。現場の教員のコメントをもとに、問題文のワ―ディン

グをその学年に適するよう修正も加えた。こうした改訂を経て、子ども特別調査学力テストの問題項目は決定された。

3　分析方法

（1）信頼性の検討方法

子ども特別調査学力テストの信頼性を検討するために、主に依拠した分析方法は、以下の 4 つである。

① テトラコリック相関係数

テストの信頼性を検討するにあたり、テトラコリック相関係数（四分相関係数）を導入した。テトラコリック相関とは、2 変数が共に順序尺度をとる離散変数であり、共にカテゴリ数を 2 つ伴う時に､最尤推定法により両変数間に求められる 2×2 の相関である（豊田, 1998； Greene and Hensher, 2010)。このテトラコリック相関では、2 値の背後に正規分布する潜在変数を仮定し、2 つの反応カテゴリ間の距離を閾値で調整する。たとえばある疾患に対する罹患群と健常群の 2 カテゴリを伴う変数間、あるいは学力テストのような正誤データを伴う変数間の相関を記述するのに有効な方法である。

本稿では、各学年のすべてのケースが同じカテゴリに範疇化された分散の無い項目（全員が正解、あるいは不正解）を除外し、かつ、すべてのケースが 2 項目以上に同じ反応カテゴリを示し、項目間の相関が 1 となる場合には、1 項目を残し他の項目を除外した後、各学年全項目を用いて、テトラコリック相関行列を算出した。

② カテゴリカル因子分析

カテゴリカル因子分析は、構造方程式モデリングの下位モデルの一つである（豊田, 1998)。2 値データを解析するための 1 因子のカテゴリカル因子分析は、以下の式で表現される。

$$z_{ij} = \alpha_j f_i + e_{ij}$$

この時、α_j は項目 $_j$ に対する因子パタン（因子負荷)、f_i は協力者 $_i$ の共通因子、e_{ij} は項目 $_j$ の協力者 $_i$ の誤差因子である。z_{ij} は、通常の因子分析では観測変数であるが、カテゴリカル因子分析では潜在変数となる。γ_j を項目 $_j$ の閾値とし、実際に観測される 0 か 1 の 2 値変数を u_{ij} としたとき、$z_{ij} > \gamma_j$ であれば $u_{ij} = 1$ が、$z_{ij} < \gamma_j$ であれば $u_{ij} = 0$ が観測される。最尤法を用いて推定されるパラメータは、因子パタン α_j 及び閾値 $\gamma_j (j = 1, \cdots, n)$ の 2 種類である。構造方程式モデリングによる 2 値データのカテゴリカル因子分析は、2 母数の項目反応理論モデルに相当することが知られている（豊田, 2003)。

本稿では、各項目の因子パタン及び閾値を推定するにあたり、算数/数学と国語に関しては、学年ごとに問題が異なるため、各学年各項目について求めたが、推論に関しては、小

学 1 年~小学 3 年、小学 4 年~小学 6 年、中学 1 年~中学 3 年で問題が共通であるため、3 学年分のデータを合わせて各項目について求めた。

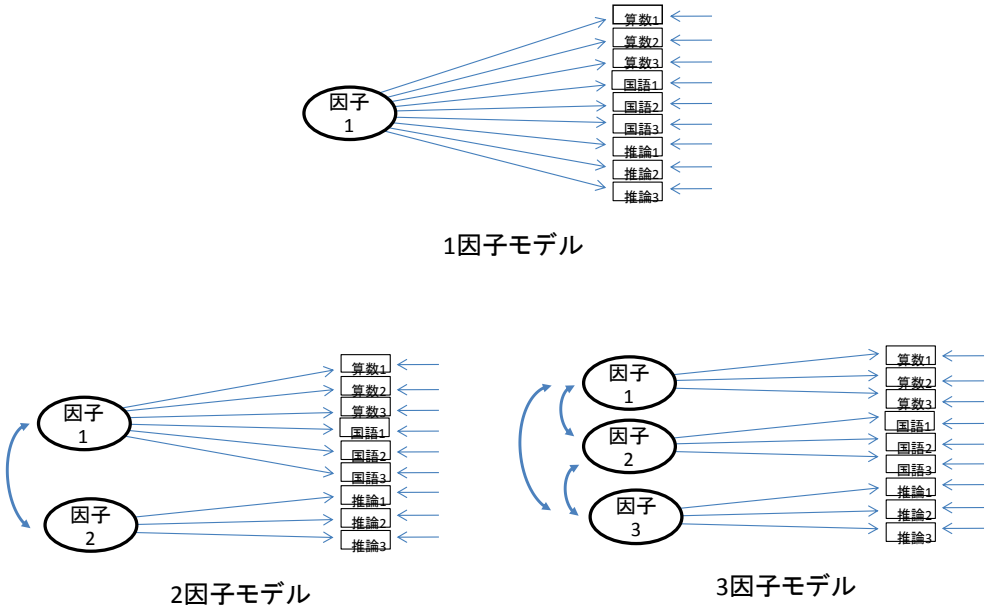
③ 確証的因子分析

母数に関する事前の知識をモデルに組み込み、仮説的構造を仮定し、実質科学的な知見からモデルを特定して行う因子分析が、確証的因子分析（confirmatory factor analysis）である（柳井他, 1990；豊田, 2000）。つまり確証的因子分析とは、最初から因子構造を定めた上でモデルの適合を確認する方法であり、探索的に因子を求める通常の（探索的）因子分析モデルの一部の影響指標を、理論的に予測される値に固定したものと考えることができる。

本稿では、学年ごとに、算数/数学、国語、推論の項目群について、i）すべての項目を 1 次元で説明する 1 因子モデル、ii）算数/数学と国語の項目を合わせて 1 次元、推論の項目を相関のある別の 1 次元で説明する 2 因子モデル、iii）算数/数学、国語、推論それぞれの項目を、相関し合う 3 次元で説明する 3 因子モデルの間で、モデルの比較を行うことにより、項目群の構造を検討した（図３－１）。

モデルの比較は、モデル適合に基づき、モデルの評価には RMSEA（root mean square error of approximation）、CFI（comparative fit index）による比較、並びにカイ 2 乗検定を併用した。RMSEA は、モデルの分布と真の分布との乖離を 1 自由度当たりの量として表現した指標であり、0.05 以下であれば、モデルの当てはまりがよく、0.1 以上であれば、当てはまりが悪いと判断される（豊田, 1998）。CFI は、観測変数間に相関が無いことを仮

図３－１　３つの確証的因子分析モデル



定した独立モデルと比較して、モデルの適合度がどの程度改善されたかで評価される。値は 0.0 から 1.0 をとり、1.0 に近いほどモデルの適合度は高いと判定される。

④ Cronbach α 係数

テストの信頼性の検定には、とりわけ尺度の項目群の内的一貫性を推定する信頼性係数の指標として、最も頻繁に用いられる Cronbach α 係数（Cronbach,1951)を用いた。

Cronbach α 係数は、項目数を N、項目間の相関係数の平均を$\bar{p}$とした時、以下の式で求められる。

$$\alpha = \frac{N\bar{p}}{1+\bar{p}(N-1)}$$

α 係数が 1 に近ければ、テストにおける回答の安定性、一貫性は高いとみなされ、そのテストの信頼性は高いと評価できるが、通常、0.8 程度ならば十分な一貫性があり、0.7 程度でもある程度は一貫性があると考えられている。また、一般に、項目数が増えると α 係数は高くなることが知られている（Nunnaly, 1978)。

本稿では、各学年の項目群から構成されたテトラコリック相関行列より、相関係数の平均を算出し、因子ごとに Cronbach α 係数を求めた。算数/数学と国語については、それぞれ学年ごとに求めたが 、推論については、小学 1 年~小学 3 年、小学 4 年~小学 6 年、中学 1 年~中学 3 年で問題が共通であるため、3 学年分のデータを用いて算出した。

（2）妥当性の検討方法

子ども特別調査学力テストのテスト項目の妥当性を検討するために、まず、子ども特別調査学力テストと予備調査の間､そして､子ども特別調査と都道府県別学力テストの間で、それぞれ対応のある項目について、正答率の平均値の比較分析を行った。

子ども特別調査では、予備調査の結果を反映させて項目を追加したため、一部の問題は予備調査に含まれていない（小学 1 年の国語 5 項目、小学 2 年の算数 1 項目、小学 3 年の算数 2 項目、中学 1 年の数学 1 項目、中学 2 年の数学 2 項目と国語 4 項目、中学 3 年の数学 4 項目と国語 4 項目、小学 4 年～中学 3 年の推論 1 項目: 表３－９を参照のこと)。これらの項目以外の共通する項目について、それぞれの正答率をケースとし、両調査間で対応のあるサンプルの t 検定を行った。

子ども特別調査と都道府県別学力テストについても同様の検定を行った。都道府県別学力テストと共通するのは、小学 4 年から中学 3 年までの算数/数学、国語の問題である（小学 6 年の算数 1 項目、中学 2 年の数学 1 項目と国語 2 項目、中学 3 年の数学 4 項目、国語 2 項目を除く: 表３－９を参照のこと)。子ども特別調査学力テストの問題の出所となった秋田県、新潟県、岐阜県、香川県が、それぞれ公表している、子ども特別調査との間で対応のある都道府県別学力テスト項目の正答率を引用し、両調査間で正答率の平均値を比較

した。都道府県別学力テストの学年上のテスト実施時期が、子ども特別調査実施時期と比べ 1 年以上離れているケース（19 項目)、及び問題を改変して比較不可能と判断されたケース（6 項目）については分析から除外した（表３－９を参照のこと)。

続いて、子ども特別調査と予備調査、子ども特別調査と都道府県別学力テストの間で、上述の対応のある項目の正答率をそれぞれケースとした時の、両調査間の相関関係を、Pearson の積率相関係数を用いて求めた。都道府県別学力テストについては、４つの県それぞれについても、平均値の比較、相関係数の算出を行った。

算数/数学、国語のテスト項目で測定される構成概念の妥当性を検討するためには、同じ構成概念を測定する別の指標として、回答時の子どもの学校での算数/数学と国語の成績を用いた。子どもの成績は、親が回答するアンケート調査の中で、それぞれ「上位」から「下位」までの 5 段階スケールで尋ねられた。各学年の算数/数学ならびに国語の因子得点と、この親が報告する 2 科目の成績との間の相関の程度を Pearson の積率相関係数を用いて求めた。

テトラコリック相関係数と、各項目の閾値と因子パタンの算出、確証的カテゴリカル因子分析には、構造方程式モデリングに適した統計ソフト Mplus 6.1 を、その他の分析には SPSS18.0 を用いた。

第３節　結果

１　テスト項目の因子構造と内的一貫性

カテゴリカル因子分析を構造方程式モデリングの確証的因子分析の枠組みで行い、小学 1 年から中学 3 年までの 9 学年について、それぞれのテスト項目を 1 因子モデル、2 因子モデル、3 因子モデルで適合した結果を表３－３に掲げる。

小学 2 年、小学 6 年、中学 1 年、中学 2 年、中学 3 年については、RMSEA の値が 3 因子モデルにおいて最も小さく、3 因子モデルの当てはまりが、2 因子モデル、1 因子モデルと比較して、よりよいことが明らかにされた（表３－３)。小学 1 年、小学 3 年、小学 4 年、小学 5 年では、3 因子モデルと 2 因子モデルの RMSEA 値は同等であり、カイ 2 乗検定も 2 因子モデルが 3 因子モデルに比べて、モデルの当てはまりを有意に悪くしないことを示したが、CFI の値は、3 学年とも 3 因子モデルの方が若干優れていた。小学 1 年、小学 2 年、小学 3 年、小学 5 年、中学 3 年においては、RMSEA が 0.03 以下、CFI も 0.97 以上と非常に優れ、モデルの当てはまりが極めてよいことが示された。

推定された各項目の因子パタン及び閾値を表３－４と表３－５に示す。因子パタンが、値は小さいが負になるケースが中学 2 年と中学 3 年の国語に 1 項目ずつ見られたが、項目への負荷は概ね高く、同等であった。

学年ごとに求めた因子間相関係数を表３－６に記した。相関係数のレンジは幅広いが、ほとんどの学年において、値は概して高かった。とくに、「算数/数学」と「国語」の 2 つ

の構成概念間の相関は、例外的に有意な相関を示さなかった中学 1 年と、相関は有意だが低かった中学 2 年を除き、0.76 以上と非常に高かった。「算数/数学」と「推論」、「国語」と「推論」の相関は、0.52 以下であり、「算数/数学」と「国語」の間の相関ほど高くはなかった。

表３－３　確証的因子分析モデル適合

		小1	小2	小3	小4	小5	小6	中1	中2	中3
3因子モデル	RMSEA	**0.02**	**0.01**	**0.01**	**0.04**	**0.02**	**0.07**	**0.06**	**0.06**	**0.03**
	CFI	**0.98**	**0.97**	**0.99**	**0.87**	**0.98**	**0.88**	**0.72**	**0.72**	**0.99**
2因子モデル	RMSEA	**0.02**	0.02	**0.01**	**0.04**	**0.02**	0.08	0.08	0.07	0.09
	CFI	**0.97**	0.90	**0.98**	**0.86**	**0.97**	0.87	0.50	0.62	0.86
	⊿χ 2	**2.13**	9.51	**4.79**	**3.76**	**3.14**	7.00	52.87	22.67	122.09
	⊿df	**2**	2	**2**	**2**	**2**	2	2	2	2
	p	***NS***	***	******	***NS***	***NS***	**	***	***	***
1因子モデル	RMSEA	0.04	0.03	0.01	0.05	0.04	0.10	0.08	0.07	0.09
	CFI	0.91	0.89	0.96	0.79	0.89	0.78	0.50	0.60	0.86
	⊿χ 2	32.12	11.89	8.58	18.79	24.08	56.15	55.23	28.18	122.47
	⊿df	3	3	3	3	3	3	3	3	3
	p	***	***	**	***	***	***	***	***	***

太字は最適モデルを示す。
***、**、*は係数がそれぞれ1%、5%、10%の水準で統計的に有意なことを示す。

表３－４　算数/数学・国語各項目の因子パタンと閾値

	小学1年			小学2年			小学3年			小学4年			小学5年			小学6年			中学1年			中学2年			中学3年		
	設問	因子パタン	閾値	設問	因子パタン	閾値	設問	因子パタン	閾値	設問	因子パタン	閾値	設問	因子パタン	閾値	設問	因子パタン	閾値	設問	因子パタン	閾値	設問	因子パタン	閾値	設問	因子パタン	閾値
算数/数学	1(1)	0.78	−1.40	1(1)	0.81	−1.69	1	0.95	−1.86	1(1)	0.88	−0.80	1(1)	0.74	−2.14	1(1)	0.77	−1.12	1(1)	0.87	−0.69	1(1)	0.69	−0.71	1	0.83	−0.30
	1(2)	0.69	−1.21	1(2)	0.70	−0.91	2	0.14	−1.22	1(2)	0.60	−0.73	1(2)	0.19	−1.21	1(2)	0.76	−1.12	1(2)	0.63	−0.93	1(2)	0.78	−1.23	2(1)	0.68	−0.18
	2(1)	0.88	−1.85	2(1)	−	−	3(1)	0.15	−1.22	1(3)	0.65	−1.72	1(3)	0.78	−0.60	1(3)	0.81	−0.56	1(3)	0.95	−0.93	1(3)	0.71	−0.58	2(2)	0.91	0.12
	2(2)	0.64	−1.30	2(2)	−	−	3(2)	0.28	−1.14	1(4)	0.64	−1.37	1(4)	0.80	−0.55	1(4)	0.73	0.81	1(4)	0.83	−0.58	1(4)	0.79	−0.33	3(1)	0.91	−0.57
	3(1)	0.60	−1.40	3(1)	0.45	−1.21	4	0.67	−1.14	2(1)	0.42	−0.53	2(1)	0.24	−0.70	2	0.80	−0.34	2	0.58	−0.63	2(1)	0.58	−0.11	3(2)	0.81	−0.06
	3(2)	0.52	−0.29	3(2)	0.61	−0.83	5(1)	0.70	−1.53	2(2)	0.59	0.24	2(2)	0.78	−0.81	3(1)	0.92	−0.20	3(1)	0.74	−0.63	2(2)	0.72	0.06	3(3)	0.81	−1.18
	4(1)	0.85	−1.21	4	0.53	−1.10	5(2)	0.30	0.10	2(3)	0.58	−1.25	3	0.68	−0.04	3(2)	0.73	−0.56	3(2)	0.66	−0.29	3(1)	0.46	−0.51	4(1)	0.85	−0.12
	4(2)	0.67	−0.81	5	0.68	−1.21	6_1	0.42	−0.39	2(4)	0.89	−0.95	4(1)	0.82	−0.65	4(1)	0.76	0.41	3(3)	0.42	0.38	3(2)	0.81	−0.71	4(2)	0.98	−0.37
	5(1)	0.75	−2.14	6	0.62	−1.34	6_2	−	−	3(1)	0.45	−0.53	4(2)	0.81	−0.99	4(2)	0.90	0.13	4(1)	0.42	−0.20	4(1)	0.86	−0.86	5(1)	1.00	−0.43
	5(2)	0.68	−1.52	7(1)	0.92	−2.00	7	0.62	−1.86	3(2)	0.93	−1.14				4(3)	0.43	−0.63	4(2)	0.40	0.38	4(2)	0.41	−0.06	5(2)	0.85	0.00
	5(3)	0.79	−1.40	7(2)	0.92	−2.00	8	0.40	−1.41																		
	5(4)	0.60	−1.06	7(3)	0.91	−1.69	9(1)	0.95	−1.86																		
	6(1)	0.49	−0.75	7(4)	0.92	−2.00	9(2)	0.95	−1.86																		
	6(2)	0.49	−0.60	8(1)	0.16	−1.34	9(3)	0.36	−1.41																		
				8(2)	0.40	−1.21	9(4)	0.45	−1.14																		
				9_1	0.03	−1.21	9(5)	0.67	−1.31																		
				9_2	0.43	−0.41	9(6)	0.86	−1.22																		
							10(1)	0.88	−1.67																		
							10(2)	0.52	−0.77																		
国語	7(1)	0.91	−1.30	10(1)	0.51	−0.67	11(1)	0.69	−1.86	4(1)	0.73	−2.03	5(1)	0.64	−1.66	5(1)	0.84	−0.72	5(1)	0.74	−1.08	5(1)	0.39	−0.64	6(1)	0.71	−1.07
	7(2)	0.96	−1.21	10(2)	0.25	−1.10	11(2)	0.83	−1.53	4(2)	0.73	−2.03	5(2)	0.84	−0.75	5(2)	0.78	−1.00	5(2)	0.46	−1.36	5(2)	0.27	−1.71	6(2)	0.77	−0.88
	7(3)	0.91	−1.52	10(3)	−	−	11(3)	0.90	−2.15	4(3)	0.95	−1.04	5(3)	0.52	0.16	5(3)	0.84	−0.56	5(3)	0.90	−0.93	5(3)	0.63	−0.51	6(3)	0.33	−1.98
	8(1)	−	−	11(1)	0.57	−1.00	11(4)	0.75	−1.07	5(1)	0.55	−1.25	6(1)	0.69	0.16	5(4)	0.45	−0.27	5(4)	0.65	−0.16	6(1)	0.73	−1.03	7(1)	0.36	−0.64
	8(2)	−	−	11(2)	0.58	−0.83	12(1)	0.25	−0.82	5(2)	0.79	−0.95	6(2)	0.54	−0.70	6(1)	0.77	−0.34	6(1)	0.76	−0.87	6(2)	0.68	−0.86	7(2)	0.71	−0.79
	8(3)	0.85	−1.85	12(1)	0.69	−1.00	12(2)	0.63	−0.57	5(3)	0.72	−0.08	6(3)	0.87	−1.52	6(2)	0.63	−0.41	6(2)	0.79	−1.25	6(3)	0.69	0.06	7(3)	0.82	−1.31
	8(4)	0.95	−1.66	12(2)	0.77	−1.69	12(3)	0.73	−1.00	5(4)	0.49	−0.19	6(4)	0.90	−1.30	6(3)	0.89	−0.56	6(3)	0.94	−1.47	6(4)	0.39	−1.71	7(4)	0.63	−1.47
	8(5)	0.46	−0.99	12(3)	0.92	−1.69	12(4)	0.49	−0.35	5(5)	0.90	−0.35	6(5)	0.89	−1.40	6(4)	0.88	0.20	6(4)	0.71	−0.93	6(5)	0.63	0.33	7(5)	0.72	0.37
	9_1	0.93	−1.66	12(4)	0.86	−0.91	12(5)	0.21	−1.00	6	0.71	−1.25	7(1)	0.69	−2.14	6(5)	0.85	−0.20	6(5)	0.64	−0.02	7	−0.10	−1.03	8(1)	−0.02	−1.07
	9_2	−	−	12(5)	0.62	−1.49	12(6)	0.34	−0.06	7	0.69	−0.87	7(2)	0.43	−0.93	7(1)	0.24	−0.81	7	0.62	−1.16	8(1)	0.63	1.23	8(2)	0.75	−0.79
	9_3	−	−	13(1)			13(1)	0.46	−0.18	8	0.53	−0.35	8(1)	0.64	−1.40	7(2)	0.76	−0.34	8(1)	0.63	−1.16	8(2)	0.52	−0.28	8(3)	0.87	−0.57
	10(1)	0.90	−1.66	13(2)	0.18	0.23	13(2)	0.47	0.06	9(1)	0.37	−0.80	8(2)	0.18	−1.40	8	0.73	−0.72	8(2)	0.53	−1.25	9(1)	0.10	−0.94	8(4)	0.24	0.97
	10(2)	0.91	−1.85	13(3)			14(1)	−	−	9(2)	0.72	−2.03	8(3)	0.33	−0.29	9	0.50	−0.72	9	0.18	−0.07	9(2)	0.13	−0.58	9(1)	0.63	0.88
	10(3)	0.96	−1.66	14(1)			14(2)	−	−													9(3)	0.93	0.11	9(2)	0.49	−0.79
	11(1)	0.84	−1.85	14(2)	0.83	−0.11	15(1)	−	−													9(4)	0.28	1.23	10	0.28	−1.18
	11(2)	−	−	14(3)			15(2)	0.94	−1.86																		
	11(3)_1	0.82	−1.66				16(1)	0.53	−1.00																		
	11(3)_2	0.72	−1.40				16(2)	0.33	−0.94																		
	11(4)_1	0.42	0.12																								
	11(4) 2	0.88	−1.85																								

表３－５　推論各項目の因子パタンと閾値

小学1年～小学3年			小学4年～小学6年			中学1年～中学3年		
設問	因子パタン	閾値	設問	因子パタン	閾値	設問	因子パタン	閾値
(1)	0.70	−0.76	(1)	0.78	−0.03	(1)	0.88	−0.29
(2)	0.87	−0.66	(2)	0.80	−0.57	(2)	0.78	−1.19
(3)	0.87	−0.72	(3)	0.85	0.06	(3)	0.44	−0.03
(4)	0.80	−0.39	(4)	0.66	−0.09	(4)	0.36	−0.17

表３－６　３因子モデルによる因子間相関（Pearson's *r*）

	小1	小2	小3	小4	小5	小6	中1	中2	中3
算数/数学—国語	.99 ***	.76 ***	.88 ***	.98 ***	.88 ***	.89 ***	.19	.36 **	.97 ***
算数/数学—推論	.34 ***	.48 ***	.27 **	.27 *	.20	−.03	.08	.24	.50 ***
国語—推論	.33 ***	.37 **	.20	.25 *	.20	.15	.15	.42 ***	.52 ***

***、**、*は係数がそれぞれ1%、5%、10%の水準で統計的に有意なことを示す。

表３－７　各学年各科目項目の内的一貫性（Cronbach α）

	算数/数学		国語		推論	
	α	項目数	α	項目数	α	項目数
小1	0.89	14	0.96	15		
小2	0.90	15	0.82	11	0.88	4
小3	0.89	19	0.85	15		
小4	0.83	10	0.88	13		
小5	0.83	9	0.87	13	0.85	4
小6	0.90	10	0.89	13		
中1	0.87	10	0.89	13		
中2	0.87	10	0.79 (0.81)	15 (14)	0.69	4
中3	0.95	10	0.85 (0.87)	15 (14)		
平均	0.88		0.87 (0.88)		0.80	

括弧内は因子パタンが負を示した項目を除外した場合

Cronbach α 係数は、算数/数学においては、最低でも小学４年と小学５年の 0.83 であり、最高は中学３年の 0.95、全学年の平均は 0.88 であり、値は十分に大きく、高い内的一貫性が確認された（表３－７）。国語においても、最低が中学２年の 0.79（因子パタンが負を示した１項目を除外すれば 0.81）、最高が小学１年の 0.96 であり、全学年の平均は 0.87 と、やはり内的一貫性は十分に高い。推論においては、中学１年から中学３年が 0.69 と低いが、小学１年から小学３年では 0.88 と高く、平均は 0.80 であった。

２　予備調査・都道県別学力テストとの正答率比較

子ども特別調査と予備調査の正答率では、小学１年から小学６年において、算数（$t = 6.37$, $df = 75$, $p < 0.001$）、国語（$t = 8.43$, $df = 87$, $p < 0.001$）、推論（$t = 6.94$, $df = 20$, $p < 0.001$）のいずれについても、子ども特別調査の方が有意に高かった（表３－８：表３－９）。中学１年から中学３年においても、算数（$t = 2.43$, $df = 24$, $p < 0.05$）、国語（$t = 2.50$, $df = 34$, $p < 0.05$）、推論（$t = 2.40$, $df = 8$, $p < 0.05$）のいずれについても、子ども特別調査の方が有意に高かった。

表３－８　子ども特別調査・予備調査正答率比較（小学１年~小学３年）

科目	小学1年			小学2年			小学3年		
	子ども特別調査		予備調査	子ども特別調査		予備調査	子ども特別調査		予備調査
	設問	正答率	正答率	設問	正答率	正答率	設問	正答率	正答率
算数	1(1)	0.93	0.94	1(1)	0.95	0.91	1	0.97	0.93
	1(2)	0.90	0.81	1(2)	0.82	0.56	2	0.89	0.73
	2(1)	0.98	0.94	2(1)	1.00	0.92	3(1)	0.89	0.85
	2(2)	0.92	0.88	2(2)	1.00	0.98	3(2)	0.87	0.88
	3(1)	0.93	0.92	3(1)	0.89	0.78	4	0.87	0.78
	3(2)	0.62	0.49	3(2)	0.82	0.48	5(1)	0.94	0.90
	4(1)	0.90	0.49	4	0.86	NA	5(2)	0.46	NA
	4(2)	0.80	0.80	5	0.89	0.95	6_1	0.65	0.63
	5(1)	0.98	0.94	6	0.91	0.77	6_2	0.60	0.49
	5(2)	0.95	0.84	7(1)	0.98	0.95	7	0.97	0.93
	5(3)	0.93	0.86	7(2)	0.98	0.95	8	0.92	0.87
	5(4)	0.87	0.68	7(3)	0.95	0.92	9(1)	0.97	NA
	6(1)	0.79	0.76	7(4)	0.98	0.92	9(2)	0.97	0.99
	6(2)	0.74	0.64	8(1)	0.91	0.81	9(3)	0.92	0.88
				8(2)	0.89	0.69	9(4)	0.87	0.82
				9_1	0.89	0.81	9(5)	0.90	0.90
				9_2	0.66	0.44	9(6)	0.89	0.91
							10(1)	0.95	0.96
							10(2)	0.78	0.75
国語	7(1)	0.92	0.82	10(1)	0.75	0.11	11(1)	0.97	0.96
	7(2)	0.90	0.86	10(2)	0.86	0.91	11(2)	0.94	0.82
	7(3)	0.95	0.82	10(3)	1.00	0.97	11(3)	0.98	0.96
	8(1)	1.00	0.96	11(1)	0.84	0.09	11(4)	0.86	0.73
	8(2)	1.00	0.94	11(2)	0.80	0.66	12(1)	0.79	0.71
	8(3)	0.98	0.82	12(1)	0.84	0.73	12(2)	0.71	0.58
	8(4)	0.97	NA	12(2)	0.95	0.88	12(3)	0.84	0.82
	8(5)	0.85	0.58	12(3)	0.95	0.78	12(4)	0.63	0.80
	9_1	0.97	0.92	12(4)	0.82	0.02	12(5)	0.84	0.83
	9_2	0.97	0.92	12(5)	0.93	0.85	12(6)	0.52	0.32
	9_3	0.97	0.86	13(1)	0.84	0.31	13(1)	0.57	0.52
	10(1)	0.97	0.94	13(2)	0.43	0.13	13(2)	0.48	0.16
	10(2)	0.98	0.92	13(3)	0.82	0.22	14(1)	1.00	0.99
	10(3)	0.97	0.86	14(1)	0.68	0.34	14(2)	1.00	0.97
	11(1)	0.98	0.98	14(2)	0.55	0.17	15(1)	1.00	0.97
	11(2)	1.00	0.98	14(3)	0.75	0.28	15(2)	0.97	0.96
	11(3)_1	0.97	NA				16(1)	0.84	0.73
	11(3)_2	0.93	NA				16(2)	0.83	0.50
	11(4)_1	0.46	NA						
	11(4)_2	0.98	NA						
推論	12(1)	0.67	0.56	15(1)	0.82	0.63	17(1)	0.86	0.64
	12(2)	0.62	0.56	15(2)	0.77	0.64	17(2)	0.86	0.67
	12(3)	0.66	0.54	15(3)	0.84	0.61	17(3)	0.83	0.76
	12(4)	0.49	0.22	15(4)	0.82	0.50	17(4)	0.70	0.62

一方、子ども特別調査と都道府県別学力テストでは、小学 4 年から小学 6 年において、算数（$t = -2.27$, $df = 22$, $p < 0.05$）については、子ども特別調査の方が有意に正答率が低かったが、国語（$t = -0.84$, $df = 34$, $p = 0.41$）には、有意差が無かった。中学１年から中学３年においては、算数に有意差は無かったが（$t = 0.95$, $df = 14$, $p = 0.36$)、国語は子ども特別調査の方が有意に高かった（$t = 3.06$, $df = 30$, $p < 0.01$)。

小中学校全学年を合わせ、全科目から対応のある項目の正答率を比較しても、子ども特別調査と予備調査では、子ども特別調査の方が有意に正答率が高かったが（$t = 11.79$, $df = 253$, $p < 0.001$)、子ども特別調査と都道府県別学力テストでは、有意差は無かった（$t = 0.95$, $df = 103$, $p = 0.34$)。

続く、対応のある項目の正答率の相関分析は、子ども特別調査と予備調査との間で、小学１年~小学３年、中学１年~中学３年については、３科目とも 0.8 程度の相関を示したが、小学４年~小学６年の算数においては、0.55 と低めの値を示した（表３－１０)。予備調査実施時において、小学５年の分数の演算が未習であったことが、協力校の教員のコメントから明らかにされている。該当する２項目（表３－９の小学５年の算数 1(1)、3）を除外し

表３－９　子ども特別調査・予備調査・都道府県別学力テスト正答率比較（小学４年~中学３年）

小学4年

科目	子ども特別調査 設問	子ども特別調査 正答率	予備調査 正答率	都道府県別テスト 正答率	学年	実施年月		県
算数	1(1)	0.79	0.68	0.87	小4	H18	7	秋田
	1(2)	0.77	0.51	0.80	小4	H18	7	秋田
	1(3)	0.96	0.76	0.96	小4	H18	7	秋田
	1(4)	0.91	0.87	0.97	小4	H19	7	秋田
	2(1)	0.70	0.17	0.54	小4	H18	1	新潟
	2(2)	0.40	0.13	0.46	小4	H16	1	新潟
	2(3)	0.89	0.64	0.85	小4	H18	1	新潟
	2(4)	0.83	0.45	0.79	小4	H18	1	新潟
	3(1)	0.70	0.55	0.78	小4	H16	1	新潟
	3(2)	0.87	0.51	0.95	小4	H18	1	新潟
国語	4(1)	0.98	0.89	0.97	小4	H18	7	秋田
	4(2)	0.98	0.93	0.98	小4	H16	1	新潟
	4(3)	0.85	0.45	0.87	小4	H17	7	秋田
	5(1)	0.89	0.66	0.89	小4	H19	7	秋田
	5(2)	0.83	0.77	0.81	小4	H18	7	秋田
	5(3)	0.53	0.02	0.75	小4	H20	12	秋田
	5(4)	0.57	0.25	0.72	小4	H20	12	秋田
	5(5)	0.64	0.25	0.61	小5	H20	1	岐阜
	6	0.89	0.66	0.88	小4	H20	12	秋田
	7	0.81	0.19	0.78	小4	H17	7	秋田
	8	0.64	0.34	0.69	小4	H19	7	秋田
	9(1)	0.79	0.72	0.87	小5	H18	4～5	香川
	9(2)	0.98	0.91	0.90	小5	H18	4～5	香川
推論	10(1)	0.40	0.38	NA				
	10(2)	0.62	0.59	NA				
	10(3)	0.40	NA	NA				
	10(4)	0.45	0.34	NA				

小学5年

科目	子ども特別調査 設問	子ども特別調査 正答率	予備調査 正答率	都道府県別テスト 正答率	学年	実施年月		県
算数	1(1)	0.98	0.15	0.95	小5	H20	1	岐阜
	1(2)	0.89	0.77	0.96	小4	H18	7 *	秋田
	1(3)	0.73	0.35	0.88	小5	H18	7	秋田
	1(4)	0.71	0.68	0.69	小5	H15	2	岐阜
	2(1)	0.76	0.74	0.79	小4	H18	1 *	新潟
	2(2)	0.79	0.47	0.74	小5	H16	1	新潟
	3	0.52	0.06	0.54	小5	H19	7	秋田
	4(1)	0.74	0.32	0.63	小5	H20	12 **	秋田
	4(2)	0.84	0.79	0.87	小5	H16	1	新潟
国語	5(1)	0.95	0.85	0.87	小4	H17	7 *	秋田
	5(2)	0.77	0.53	0.63	小5	H21	4～5	香川
	5(3)	0.44	0.12	0.26	小5	H21	4～5	香川
	6(1)	0.44	0.15	0.38	小5	H20	1	岐阜
	6(2)	0.76	0.91	0.62	小5	H17	1	岐阜
	6(3)	0.94	0.82	0.92	小5	H16	1	岐阜
	6(4)	0.90	0.65	0.86	小5	H18	1	新潟
	6(5)	0.92	0.38	0.81	小4	H18	7 *	秋田
	7(1)	0.98	0.85	0.94	小5	H16	1	新潟
	7(2)	0.82	0.62	0.75	小5	H16	1	新潟
	8(1)	0.92	0.97	0.87	小5	H18	4～5	香川
	8(2)	0.92	0.85	0.90	小5	H18	4～5	香川
	8(3)	0.61	0.35	0.72	小6	H20	4～5	香川
推論	9(1)	0.55	0.50	NA				
	9(2)	0.77	0.71	NA				
	9(3)	0.52	NA	NA				
	9(4)	0.55	0.50	NA				

小学6年

科目	子ども特別調査 設問	子ども特別調査 正答率	予備調査 正答率	都道府県別テスト 正答率	学年	実施年月		県
算数/数学	1(1)	0.87	0.85	0.95	小5	H20	1 *	岐阜
	1(2)	0.87	0.82	0.83	小6	H15	2	岐阜
	1(3)	0.71	0.90	0.90	小6	H15	2	岐阜
	1(4)	0.21	0.11	0.38	小6	H20	4～5	香川
	2	0.63	0.66	0.74	小6	H18	1	新潟
	3(1)	0.58	0.61	0.54	小5	H19	7 *	秋田
	3(2)	0.71	0.72	0.74	小6	H18	4～5	香川
	4(1)	0.34	0.36	NA				
	4(2)	0.45	0.28	0.60	小6	H19	7	秋田
	4(3)	0.74	0.77	0.80	中1	H18	4～5	香川
国語	5(1)	0.76	0.56	0.63	小5	H21	4～5 *	香川
	5(2)	0.84	0.92	0.90	中1	H18	1	岐阜
	5(3)	0.71	0.41	0.84	小6	H20	12	秋田
	5(4)	0.61	0.38	0.64	小6	H20	12	秋田
	6(1)	0.63	0.08	0.66	小6	H18	4～5	香川
	6(2)	0.66	0.79	0.62	小5	H17	1 *	岐阜
	6(3)	0.71	0.69	0.77	小6	H17	1	岐阜
	6(4)	0.42	0.25	0.56	小6	H20	1	岐阜
	6(5)	0.58	0.41	0.79	小6	H20	12	秋田
	7(1)	0.79	0.51	0.72	小6	H20	4～5	香川
	7(2)	0.63	0.64	0.69	小6	H20	4～5	香川
	8	0.76	0.85	0.84	小6	H20	4～5	香川
	9	0.76	0.39	0.82	小6	H18	1	岐阜
推論	10(1)	0.58	0.51	NA				
	10(2)	0.74	0.66	NA				
	10(3)	0.50	NA	NA				
	10(4)	0.63	0.46	NA				

中学1年

科目	子ども特別調査 設問	子ども特別調査 正答率	予備調査 正答率	都道府県別テスト 正答率	学年	実施年月		県
算数/数学	1(1)	0.75	NA	0.64	中1	H19	1	岐阜
	1(2)	0.83	0.85	0.91	小6	H16	1 *	岐阜
	1(3)	0.83	0.77	0.71	中1	H19	1	岐阜
	1(4)	0.72	0.65	0.60	中1	H19	1	岐阜
	2	0.74	0.77	0.74	小6	H18	1 *	新潟
	3(1)	0.74	0.58	0.76	中1	H15	2	岐阜
	3(2)	0.61	0.58	0.49	中1	H16	1	岐阜
	3(3)	0.35	0.42	0.40	中1	H20	12	秋田
	4(1)	0.58	0.46	0.73	中1	H19	7 **	秋田
	4(2)	0.35	0.23	0.73	中2	H20	4～5**	香川
国語	5(1)	0.86	0.54	0.77	中2	H18	1	新潟
	5(2)	0.91	0.89	0.90	中1	H18	1	岐阜
	5(3)	0.83	0.85	0.60	中2	H20	1	岐阜
	5(4)	0.56	0.19	0.49	中1	H19	7	秋田
	6(1)	0.81	0.69	0.77	小6	H20	12 *	秋田
	6(2)	0.90	0.69	0.87	中1	H17	1	岐阜
	6(3)	0.93	0.77	0.59	中1	H15	2	岐阜
	6(4)	0.83	0.54	0.73	中1	H20	1	岐阜
	6(5)	0.51	0.35	0.36	中1	H21	4～5	香川
	7	0.88	0.81	0.84	小6	H20	4～5*	香川
	8(1)	0.88	0.77	0.79	中1	H15	2	岐阜
	8(2)	0.90	0.73	0.69	中1	H17	7	秋田
	9	0.53	0.58	0.31	中1	H21	4～5	香川
推論	10(1)	0.61	0.42	NA				
	10(2)	0.84	0.77	NA				
	10(3)	0.58	NA	NA				
	10(4)	0.61	0.27	NA				

中学2年

科目	子ども特別調査 設問	子ども特別調査 正答率	予備調査 正答率	都道府県別テスト 正答率	学年	実施年月		県
数学	1(1)	0.76	0.72	0.60	中1	H19	1 *	岐阜
	1(2)	0.89	0.84	0.88	中2	H15	2	岐阜
	1(3)	0.72	0.72	0.61	中2	H16	1	岐阜
	1(4)	0.63	0.53	0.50	中2	H15	2	岐阜
	2(1)	0.54	NA	0.48	中3	H18	7	秋田
	2(2)	0.48	0.25	0.40	中1	H20	12 *	秋田
	3(1)	0.70	NA	NA	中3	H21	東京都立高入試	
	3(2)	0.76	0.50	0.78	中3	H19	7	秋田
	4(1)	0.80	0.66	0.75	中2	H19	1 **	岐阜
	4(2)	0.52	0.41	0.56	中1	H17	7 *	秋田
国語	5(1)	0.74	0.88	0.45	中2	H15	2	岐阜
	5(2)	0.96	0.91	0.90	中2	H15	2	岐阜
	5(3)	0.70	0.66	0.66	中2	H16	1	岐阜
	6(1)	0.85	0.75	0.66	中2	H16	1	岐阜
	6(2)	0.80	0.75	0.76	中2	H17	1	岐阜
	6(3)	0.48	0.53	0.33	中2	H18	1	岐阜
	6(4)	0.96	0.88	0.89	中2	H18	1	新潟
	6(5)	0.37	NA	0.52	中3	H19	7	秋田
	7	0.85	0.72	0.82	中2	H18	1	岐阜
	8(1)	0.11	NA	NA				
	8(2)	0.61	NA	NA				
	9(1)	0.83	0.84	0.79	中2	H18	1月	新潟
	9(2)	0.72	0.91	0.79	中1	H15	2月 *	岐阜
	9(3)	0.46	NA	0.46	中2	H15	2月	岐阜
	9(4)	0.11	0.03	0.25	中2	H19	1月	岐阜
推論	10(1)	0.65	0.59	NA				
	10(2)	0.91	0.72	NA				
	10(3)	0.39	NA	NA				
	10(4)	0.54	0.56	NA				

中学3年

	子ども特別調査 設問	子ども特別調査 正答率	予備調査 正答率	都道府県別テスト 正答率	学年	実施年月		県
u1	1	0.62	NA	NA	中3	H21	東京都立高入試	
u2	2(1)	0.57	NA	NA	中3	H20	秋田県立高入試	
u3	2(2)	0.45	NA	NA	中3	H17	愛媛県立高入試	
u4	3(1)	0.71	0.63	0.73	中3	H19	7	秋田
u5	3(2)	0.52	0.75	0.67	高1	H18	4～5	香川
u6	3(3)	0.88	0.81	0.94	高1	H18	4～5	香川
u7	4(1)	0.55	0.56	0.69	高1	H18	4～5	香川
u8	4(2)	0.64	NA	NA	中3	H20	秋田県立高入試	
u9	5(1)	0.67	0.66	0.68	中3	H17	7 **	秋田
u10	5(2)	0.50	0.56	0.61	中3	H18	7 **	秋田
u11	6(1)	0.86	0.78	0.66	中2	H16	1 *	岐阜
u12	6(2)	0.81	0.88	0.85	中3	H17	7	秋田
u13	6(3)	0.98	0.94	0.82	中3	H17	7	秋田
u14	7(1)	0.74	0.84	0.33	中2	H18	1 *	岐阜
u15	7(2)	0.79	0.84	0.83	中3	H17	7	秋田
u16	7(3)	0.91	0.94	0.91	中3	H18	7	秋田
u17	7(4)	0.93	0.91	0.90	中3	H19	7	秋田
u18	7(5)	0.36	0.45	0.52	中3	H19	7	秋田
u19	8(1)	0.86	0.97	0.85	中3	H18	7	秋田
u20	8(2)	0.79	0.63	0.58	中3	H18	7	秋田
u21	8(3)	0.71	NA	0.46	中2	H15	2 *	岐阜
u22	8(4)	0.17	NA	0.25	中2	H19	1 *	岐阜
u23	9(1)	0.19	NA	NA				
u24	9(2)	0.79	NA	NA				
u25	10	0.88	0.88	0.91	中3	H18	7	秋田
u26	11(1)	0.57	0.50	NA				
u27	11(2)	0.91	0.78	NA				
u28	11(3)	0.55	NA	NA				
u29	11(4)	0.52	0.63	NA				

*: 都道府県別学力テスト実施日と、子ども特別調査実施日（該当学年より１学年上の４～５月）が 12 カ月以上離れているケース
**: 都道府県別学力テスト実施問題を改変したケース

て相関係数を求めると、小学４年~小学６年の算数は 0.68（27 項目）と上昇した。全学年を含めた相関は、算数/数学が 0.75（101 項目）、国語が 0.74（123 項目）、推論が 0.76（30 項目）であった。

子ども特別調査と都道府県別学力テストとの間では、中学１年~中学３年の数学の 0.79（15 項目）を除いて、3 学年で範疇化した算数/数学、国語におけるすべての相関係数が 0.80 を超え、小学４年~小学６年の算数では 0.89（23 項目）と最も高かった。全学年を含めた相関は算数/数学が 0.85（38 項目）、国語が 0.82（66 項目）であった（表３－１０）。

さらに、都道府県別学力テスト情報を県別に分類した分析は、子ども特別調査と比較して、秋田県の正答率の方が有意に高く、岐阜県の正答率の方が有意に低いが、香川県、新潟県について有意差は無いことを示した（表３－１１）。小学４年から中学３年までの対応のある項目について、子ども特別調査正答率と、都道府県別学力テストの県別正答率の散布図を図３－２に示した。これら 104 項目についての相関係数 0.82（表３－１０）は、県別に求めるとさらに上昇した（表３－１１：0.82~0.88）。

表３－１０　子ども特別調査正答率と予備調査・都道府県別学力テスト正答率の相関（Pearson's *r*）

学年	科目	予備調査 *r*		予備調査 項目数	都道府県別テスト *r*		都道府県別テスト 項目数
小学1年～小学3年	算数	0.82	***	47	NA		
	国語	0.77	***	49	NA		
	推論	0.77	***	12	NA		
小学4年～小学6年	算数	0.55	***	29	0.89	***	23
	国語	0.76	***	39	0.84	***	35
	推論	0.93	***	9	NA		
中学1年～中学3年	数学	0.80	***	25	0.79	***	15
	国語	0.81	***	35	0.83	***	31
	推論	0.68	**	9	NA		
全学年	算数/数学	0.75	***	101	0.85	***	38
	国語	0.74	***	123	0.82	***	66
	推論	0.76	***	30	NA		

***、**は係数がそれぞれ1%、5%の水準で統計的に有意なことを示す。

表３－１１　子ども特別調査と都道府県別学力テスト平均正答率の県別比較（小学４年~中学３年）

	正答率平均	標準偏差	項目数	*t*値		*df*	Pearson's *r*	
子ども特別調査	0.72	0.18	34	−2.08	**	33	0.85	***
都道府県別テスト：秋田県	0.76	0.16						
子ども特別調査	0.81	0.15	16	1.03		15	0.88	***
都道府県別テスト：新潟県	0.79	0.14						
子ども特別調査	0.74	0.18	35	3.02	***	34	0.82	***
都道府県別テスト：岐阜県	0.68	0.18						
子ども特別調査	0.68	0.19	19	−0.13		18	0.84	***
都道府県別テスト：香川県	0.68	0.21						

***、**は値がそれぞれ1%、5%の水準で統計的に有意なことを示す。

図３－２　子ども特別調査と都道府県別学力テストの項目別正答率の県別分布（小学 4 年~中学 3 年）

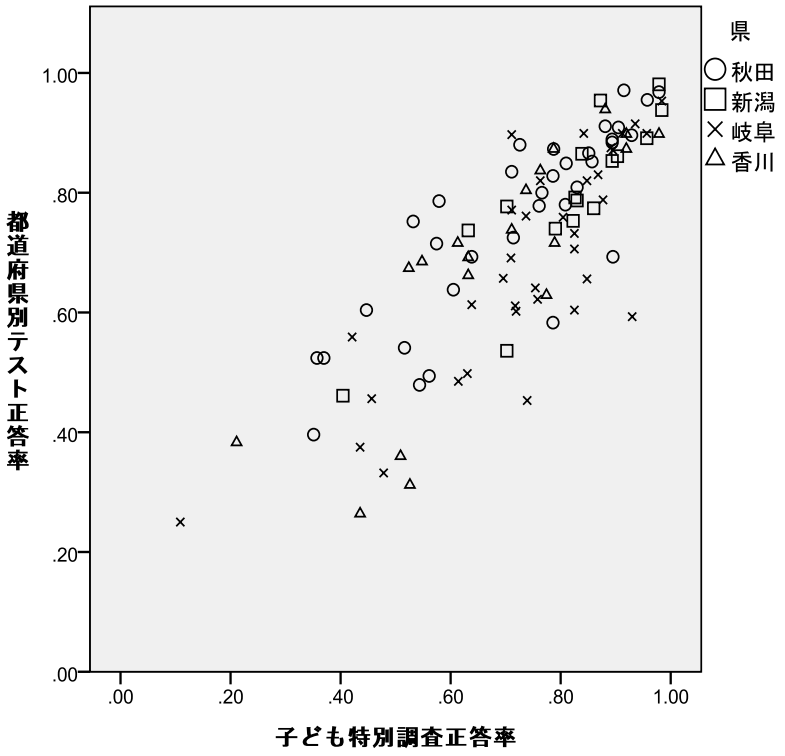


表３－１２　学校の成績と 2 科目因子得点との相関（Pearson's *r*）

		学校の成績											
		小1		小2		小3		小4		小5		小6	
		算数	国語	算数	国語	算数	国語	算数	国語	算数	国語	算数	国語
因子得点	算数	.39 ***	.44 ***	.59 ***	.47 ***	.47 ***	.41 ***	.57 ***	.54 ***	.58 ***	.44 ***	.61 ***	.57 ***
	国語	.39 ***	.44 ***	.55 ***	.43 ***	.41 ***	.39 ***	.61 ***	.57 ***	.56 ***	.49 ***	.58 ***	.62 ***

		学校の成績							
		中1		中2		中3		小1～中3平均	
		数学	国語	数学	国語	数学	国語	算数/数学	国語
因子得点	数学	.61 ***	.46 ***	.67 ***	.14	.64 ***	.52 ***	.57 ***	.44 ***
	国語	.00	.14	.52 ***	.50 **	.61 ***	.53 ***	.47 ***	.46 ***

***、**は係数がそれぞれ1%、5%の水準で統計的に有意なことを示す。

３　テスト得点と学校の成績

算数/数学の因子得点と算数/数学の成績の相関は、小学 1 年で 0.39、小学 3 年で 0.47 と低かったが、他の学年では、0.57 から 0.67 のレンジにあり、相関は高く、全学年の平均は 0.57 であった（表３－１２）。一方、国語の因子得点と国語の成績は、小学 3 年で 0.39、中学 1 年で 0.14 と低かったが、他の学年では、0.43 から 0.62 のレンジにあり、全学年の平均は 0.46 であった。

小学 1 年を除くすべての学年において、算数/数学の因子得点は、国語の成績よりも算数の成績とより高く相関している傾向を示した。一方、国語の因子得点が、算数の成績よりも国語の成績とより高く相関している傾向は、明確には見られなかった。

第４節　考察

すべての学年の算数/数学、国語、推論の各項目に施した確証的カテゴリカル因子分析は、3 因子構造の当てはまりがよいことを示し、このことは、それぞれの科目で測定される構成概念が、相関はし合うが同一ではないことを示唆している。しかしながら、いくつかの学年において、2 因子構造の当てはまりもほとんど遜色なく、算数/数学と国語を合わせて一次元で考えるモデルの妥当性も否定されない。算数/数学、国語、推論各項目の内的一貫性から評価する信頼性係数は、項目数が少ない推論で一部低かったが、概ね非常に高く、テスト項目の高い信頼性を確認することができた。

子ども特別調査学力テストの正答率と、既存の正答率データとの平均値の比較は、子ども特別調査は予備調査に比べて、中学の数学を除き、正答率が高いことを示した。しかし、予備調査データは、小学校中学校共に、静岡県の 1 校 1~2 学級から収集されたものであり、特定のサンプルとの間の正答率の比較結果が、テスト項目の構成概念妥当性を脅かすとは考えていない。むしろ、子ども特別調査が予備調査より、約 5 か月後に実施されたことを考えれば、この結果はきわめて順当であり、テスト項目の構成概念妥当性を低める証拠とはいえない。

重要なのは大規模な都道府県別学力テストとの比較であり、すべての学年のすべての項目を合わせた検定において、正答率の平均値に差が見られなかったことの意義は大きい。子ども特別調査のサンプルの代表性の検討は山下他（2011）に譲るが、子ども特別調査の標本特性が全国を代表したものであり、扱う子どもサンプルの学力水準についても、全国を代表したものであるという前提に立てば、４つの県の正答率の平均値との間に差が無かったことは、テスト項目の構成概念妥当性の高さを示唆するものと考えられる。加えて、県別の比較において、秋田県は子ども特別調査より有意に正答率が高く、岐阜県は子ども特別調査より有意に低いが、香川県と新潟県とは差が無かったことも特筆に値する。小学 6 年と中学 3 年を対象として行われる全国学力調査が開示する算数/数学と国語の正答率の都道府県別スコア（『平成 22 年度全国学力・学習状況調査・報告書・集計結果』について: 国立教育政策研究所 http://www.nier.go.jp/10chousakekkahoukoku/）は、秋田県の算数/数学と国語の学力の高さを安定して報告しており、その秋田県の正答率が、子ども特別調査より有意に高かったことは、子ども特別調査学力テスト項目の構成概念妥当性の高さを傍証するもう 1 つのエビデンスを添えるものといえるだろう。さらに、予備調査と子ども特別調査の間に対応のある全項目における相関係数は 0.75、都道府県別学力テストとは 0.82 といずれも高く、このことからも、テスト項目の構成概念妥当性の高さが確認できる。

親が報告する子どもの学校の成績と、子ども特別調査学力テスト得点との相関は、全学年の平均では、算数/数学が 0.57、国語が 0.46 であった。これらの中程度の相関は、学力テストの併存的妥当性を支持するものであると同時に、テスト得点が、親が報告する成績では説明できない成分を伴うことを示すものでもある。さらに、算数/数学のテストの方が国語より、その教科の成績との併存的妥当性がより高かったこと、加えて、算数/数学のテストには弁別的妥当性が見られたが、国語のテストに見られなかったことは、算数/数学の子ども特別調査学力テストの方が、国語に比べ、その教科の成績に対する予測力が高いことを示している。このことは、国語のテスト問題は、読解や文章記述の項目を含まず、語彙と漢字のみに特化させたのに対し、算数/数学のテスト問題はより包括的であり、学校で教科として習得している要素をより広く反映させていることによる可能性が考えられる。あるいは、教科によって、子どもの測定される学力と学校の成績評価との関連の程度が異なることを示唆するものとも解釈できる。

子ども特別調査学力テストの信頼性と妥当性は確認された。これらの検討結果は、子ども特別調査が、郵送法を用いた家庭でのセルフ・アドミニストレーションという、必ずしも精緻な測定が保証されないテスト環境のもとで実施されたものであっても、精度高く、測定すべき概念を測定していることを示唆しており、子ども特別調査学力テスト回答の有効性を裏付けるものと考えることができる。

しかしながら、本稿で対象としたのは各学年 38～63 名であり、小標本における検討であった。とくに、カテゴリカル因子分析から安定した結果を得るにはサンプルサイズが十分ではなく、学年ごとに求めた算数/数学、国語、推論の因子間の相関の程度に見られた幅が、学年の特性によるものなのか、テスト問題の違いによるものなのか、サンプルサイズの不十分さによる推定誤差から生じたものなのかも明らかではない。引き続き、サンプルサイズを拡大して検証を重ねていくことが求められる。

こうした課題は残るが、子ども特別調査学力テスト得点は、有効な学力指標変数として、JHPS の豊富なデータセットと連結した分析を行うことにより、我が国の学力と社会経済的背景変数との関連の解明に向けて、大きな進展をもたらし得るものと期待できる。

参考文献

Bara, B. G., M. Bucciarelli and P. N. Johnson-Laird (1995) “The development of syllogistic reasoning,” *American Journal of Psychology*, vol.108, pp157–193.

Cronbach, L. J. (1951) “Coefficient alpha and the internal structure of tests,” *Psychometrika,* vol.16, pp297-334.

Gottfredson, L. S. (1997) “Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography,” *Intelligence*, vol.24, pp.13-23.

Greene, W. H. and D. A. Hensher (2010) *Modeling Ordered Choices: A Primer*,

Cambridge University Press.

Nunnaly, J. (1978) *Psychometric theory*. New York: McGraw-Hill.

Shikishima, C., K. Hiraishi, S. Yamagata, Y. Sugimoto, R. Takemura, K. Ozaki, M. Okada, T. Toda and J. Ando (2009) "Is *g* an entity? A Japanese twin study using syllogisms and intelligence tests," *Intelligence*, vol.37, pp.256–267.

Shikishima, C., S. Yamagata, K. Hiraishi, Y. Sugimoto, K. Murayama and J. Ando (in press) "A simple syllogism-solving test: Empirical findings and implications for g research," *Intelligence*.

Snyderman, M, and S. Rothman (1988) *The IQ Controversy, the Media and Public Policy*, Transaction Books.

苅谷剛彦・志水宏吉編（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。

JHPS 子ども特別調査ワーキンググループ（2010）「日本家計パネル調査（JHPS）平成２１年度「子ども特別調査」利用者マニュアル（第１版）」パネル調査共同研究拠点パネルデータ設計・解析センター: 慶應義塾大学。

志水宏吉（2005）「学力を育てる」岩波新書。

橘木俊詔・松浦 司（2009）『学歴格差の経済学』勁草書房。

豊田秀樹（1998）『共分散構造分析―構造方程式モデリング[入門編] 』朝倉書店。

豊田秀樹（2000）『共分散構造分析―構造方程式モデリング[応用編] 』朝倉書店。

豊田秀樹編（2003）『共分散構造分析―構造方程式モデリング[技術編] 』朝倉書店。

樋口美雄・宮内 環・C. R. McKenzie・慶應義塾大学パネルデータ設計・解析センター編（2010）『貧困のダイナミズム―日本の税社会保障・雇用政策と家計行動』慶應義塾大学出版会。

耳塚寛明（2007）「だれが学力を獲得するのか」耳塚寛明・牧野カツコ編『学力とトランジッションの危機―閉ざされた大人への道』金子書房、pp. 3-24。

柳井晴夫・繁桝算男・前川眞一・市川雅教（1990）『因子分析―その理論と方法―』朝倉書店。

山下 絢・中村亮介・赤林英夫・直井道生・敷島千鶴（2011）「『JHPS お子様に関する特別調査』における家計の属性」パネル調査共同研究拠点 Discussion Paper。