

## 科学研究費助成事業（科学研究費補助金）研究成果報告書

平成25年 4月20日現在

機関番号：11301

研究種目：基盤研究（B）

研究期間：2009～2012

課題番号：21330150

研究課題名（和文） 異なる方法によって得られた評価結果の対応づけに関する基礎研究

研究課題名（英文） Linking methods for the scores obtained by the different assessment procedures

研究代表者

柴山 直（SHIBAYAMA TADASHI）

東北大学・大学院教育学研究科・教授

研究者番号：70240752

研究成果の概要（和文）：

この研究の目的は教育場面におけるさまざまな状況の中で、異なる方法によって得られた評価結果を互いに対応づけるいくつかの方法を開発することにある。具体的には、異なる年度に実施された学力調査のリンキング、異なる主催団体によって実施された学力調査のリンキング、法学教育における評価分析への適用、開発方法の評価のためのシミュレーション研究、学士課程教育への適用、小論文の自動採点の試みなどの場面をとりあげ、それぞれにふさわしい諸手法を開発した。

研究成果の概要（英文）：

This research developed procedures for linking methods for the scores obtained with different assessment tools in different assessment fields. It consists of many approaches as follows; (a) a linking method for achievement tests which are administered in different forms, (b) an evaluation study of standard error of equating by bootstrap method in equating test consisting of subtests, (c) the development of a computerized adaptive test system for undergraduate program in the case that item parameters are not able to be estimated exactly, (d) a case study of argument analysis system in automated essay scoring.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2009年度	4,000,000	1,200,000	5,200,000
2010年度	2,600,000	780,000	3,380,000
2011年度	2,200,000	660,000	2,860,000
2012年度	1,900,000	570,000	2,470,000
年度			
総計	10,700,000	3,210,000	13,910,000

研究分野：社会科学

科研費の分科・細目：心理学・教育心理学

キーワード：教育評価、等化、リンキング、対応づけ、学力調査、小論文、自動採点、CAT

## 1. 研究開始当初の背景

学力評価のグローバル化が本研究の背景に。PISAやTIMSSに代表されるように学力問題は一国に閉じたものではなく、自国の

独自性を保ちつつも、他国との共通の土俵の上で比較検討されるようになってきている。経済のグローバル化に伴う人材の流動化とその人材の「質」保証がその裏側にあるとい

えるであろう。その動きは、欧州においては CEFR (Common European Framework of Reference for Languages: 外国語能力水準共通参照枠) の設定になってあらわれている。また、米国においては全米学力進捗度調査 (NEAP) と TIMSS あるいは、ACT と ITED (Iowa Tests of Educational Development) との関連づけなどの動きなども別の例である。そのいずれにおいても互いの評価結果をなんらかのプロセスによって関連づけるさまざまな工夫が始められている (Dorans 他, 2007)。

振り返って日本の視点からこの問題を眺めたとき、先に挙げた義務教育段階における学力調査や高等教育段階における質保証問題以外にも、たとえば、外国語としての日本語学習者向けの言語能力検査における「日本語能力試験」と「日本留学試験 (日本語)」の並立や、法科大学院入学のための適性試験における日弁連法務研究財団実施の「法科大学院統一適性試験」と独立行政法人大学入試センター実施の「法科大学院適性試験」の 2 つの公的試験の並列などがこの研究の学術的背景の典型例として指摘できる。

## 2. 研究の目的

以上の状況を換言すれば教育評価の様々な局面において evidence-based な判断と説明は必須であり、その判断や説明の根拠となるべき基礎データは、重層構造をなしていることが多い。たとえば、学力調査における全国レベル、各自治体レベル、さらには各学校レベルなどが指摘できる。これらは、同じ学力を調査するものであっても、互いに手段、手続きが異なり、直接的には比較が不可能なものである。

そのため、それぞれの調査結果には数多くの情報が含まれているにもかかわらず、個々別々にしか扱えず、総合的な判断のためのボトルネックとなっている。この問題は義務教育現場だけに限られたものではなく、大学における学士課程教育や大学院教育の質保証の問題にも直結するものである。すなわち独自の教育課程を維持しつつ、いかにその質を教育機関全体に共通して担保していくかが、教育評価における喫緊の課題となっているのである。

本研究の主目的は、この課題に対して、教育評価・教育測定の立場から、異なる複数の手段・手続きによって得られた複数の評価結果を互いに比較可能とするような基礎的技術の開発をめざすことにある。これにより、種々の評価結果をたがいに対応づけ、ひいてはそれらを統合的に用いる可能性を開くことができる。

## 3. 研究の方法

同一の心理学的特性 (学力、能力、適性)

あるいは同一のキー・スキルではあるものの、たがいに異なる評価観点・評価基準に基づいて得られた評価結果、あるいは、異なる設計仕様のもとで作成されたテストで得られたスコア、さらにはパフォーマンス・アセスメントのように評定者の主観が本質的に入る評定結果であっても、それらを相互に対応づける手続きを開発する。その技術的基盤は、これまで教育測定・心理測定の分野で発展してきた「テストの等化 (equating) 法」にある。しかし、等化が可能な条件は、同じ設計仕様のもとで作成されたテスト間のみであるため、本研究で扱うような緩い条件下では適用できないという欠点ないし限界があり、開発される手法はそれらを超えるものとなる。

さらに、本手法の適用場面としては、(1) 義務教育、(2) 学士課程教育、(3) 大学院入試レベルまでの教育課程、ならびに各種言語能力試験までの幅広い教育現場を対象とし、評価方法としても客観的評価のみならず、パフォーマンス・アセスメントまでをカバーする。

## 4. 研究成果

(1) 異なる年度に実施された学力調査のリンク

新潟県教育委員会・新潟市教育委員会の協力のもとに、平成 16 年度および平成 18 年に新潟県で実施された全県学力調査の小学校約 570 校の小学 5 年生約 21500 名、中学校約 250 校中学 2 年生約 21700 名の算数と数学に関するデータの提供を受けた。また、あわせて平成 16 年度ならびに平成 18 年度実施問題冊子とそれを組み替えた作成した問題冊子 2 種類の計 4 種類の問題冊子を平成 23 年 1 月中旬に実施し、中学 2 年生については 33 校 2263 名のデータを収集した。このデータを係留 (アンカー) テストとして平成 16 年度、平成 18 年度に得られた全県学力データを対法づけることを目標とした分析を試みた。これはリンクのデータ収集デザインとしては係留テストを伴う不等価グループデザインと見なすことができる。多母集団モデル、EAP, MAP, MLE のそれぞれの観点から考察した結果、安定的に推定ができるのが MAP であること、また MAP および多母集団モデルによる平成 16 年度、平成 18 年度の年度間比較の結果、学力に大きな変動が見られなかったことが実証的にしめされた。

(2) 異なる主催団体によって実施された学力調査のリンク

全国規模の学力調査データと自治体テストとのリンク可能性について実データに基づき検討した。項目が開示される我が国独特のテスト文化の中で得点分布をいかに比較するかを全国テストと自治体テストとのリンク (対応づけ) の観点から実証的に考察し論文として 公刊した。この論文は

高く評価され日本テスト学会論文賞を受賞した。

### (3) 法学教育における評価分析への適用

権威主義的パーソナリティと司法参加の観点から複数の尺度で測定されたパーソナリティスコアの対応づけにもとづく成果の上にならって、さらに量刑研究への展開を試みた。

### (4) 開発方法の評価のためのシミュレーション研究

単一グループデザインでのリンクングにおけるシミュレーションモデルの構築と誤差評価、素得点分布に基づくリンクング誤差のシミュレーション評価、それらから発展する形で、下位テストから構成されるテスト間の等化において合計と等化の順序が与える影響に関してあらたな知見が得られた。この成果は対応づけの中でもっとも前提条件が厳しい等化を法科大学院統一適性試験の実データに適用したものである。

### (5) 学士課程教育への適用

高校での科目履修の多様化や大学入試の多様化による、大学入学時点での学生の学力レベルの不透明化、不均一化の問題に関する学力の効率的な測定に資するコンピュータを利用した適応型テスト (Computerized Adaptive Test; CAT) を開発した。特徴は予備調査が十分に実施できずに、問題項目に関する情報が満足に得られない状況下において、数学を対象に人間の主観によって定められた項目難易度を基準にした独自の計算システムを開発し、試験テストの対応づけによってシステムの妥当性が検証された。

### (6) 小論文の自動採点の試み

入学試験における小論文問題の普及にとりまう、採点コスト・労力の軽減のため自動採点システムを提案するとともに、人の採点結果と自動採点システムとの対応づけを試みた。システムは、人の手による判断ルールの構築、Support Vector Machine による方法、帰納論理プログラミングによる方法の順に開発され、最後に半教師付 LDA を用いた論点抽出による手法を開発することによって、相反する論点が記述されている小論文であっても、相対的に学習のコストを軽減しながら、人間の平均的な採点者と同程度の抽出精度をもつ方法 (KUBOSAWA, S, LU, Y, OKADA, S, NITTA, K (2012)) を開発した。

### (7) その他

また、以上より派生した成果として、新潟県の学力テストの結果をバックグラウンド情報としてもちいて、一般的に数学・算数の力がどの程度であれば、たとえば、いかなる状況であっても、求積公式を適切に用いることができるかなどの、質的および量的な側面からのアプローチによるルーブリック作成手続きのプロトタイプを試作をおこない、パフ

オーマンス・アセスメントへの適用可能性を見いだした。

### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕 (計 5 件)

- ①熊谷龍一、後藤謙司、中畦菜穂子、柴山直、佐藤喜一、野口裕之、項目に関する事前情報が十分に得られない状況における適応型テストシステム開発の試み、日本テスト学会誌、査読有、9 巻、2013 年、印刷中
- ②石井秀宗、安永和央、全項目が開示されるテスト文化のもとでの得点分布の経年比較—全国テストと自治体テストのリンクング—、日本テスト学会誌、査読有、7 巻、2011 年、23—25 ページ
- ③佐藤喜一、柴山直、下位テストから構成されるテスト間の等化におけるブートストラップ法を採用した等化の標準誤差の評価、日本テスト学会誌、査読有、7 巻、2011 年、85—98 ページ
- ④佐藤喜一、柴山直、単一グループデザインでのリンクングにおけるシミュレーションモデルの構築と誤差評価の具体例、日本テスト学会誌、査読有、6 巻、2010 年、51—62 ページ

〔学会発表〕 (計 13 件)

- ①柴山直、大規模学力調査における BIB デザインと推算値導入の効果について、日本教育心理学会第 55 回総会、2013 年 8 月 17-19 日、法政大学、発表決定 (抄録集印刷中)
- ②柴山直、重複テスト分冊法による学力調査の年度間比較、日本テスト学会第 9 回大会、2011 年 9 月 11 日、岡山大学
- ③柴山直、佐藤喜一、リンクングのためのシミュレーションモデルの構築、日本テスト学会第 7 回大会、2009 年 9 月 4 日、名古屋大学
- ④佐藤喜一、柴山直、リンクングにおけるシミュレーションモデルを利用した誤差評価の試み、日本テスト学会第 7 回大会、2009 年 9 月 4 日、名古屋大学

〔図書〕 (計 3 件)

- ①柴山直、佐藤喜一、等パーセントイル等化法における誤差評価の試み、商事法務、「JLF 叢書 Vol.16」、2010 年、56-72 ページ

〔その他〕

受賞

- ①日本テスト学会論文賞第 5 回 (2011 年)  
「全項目が開示されるテスト文化のもとでの得点分布の経年比較—全国テストと

自治体テストのリンキングー」(著者 石井秀宗、安永和央)

- ②日本テスト学会大会発表賞第4回(2010年)  
「下位テストから構成されるテスト間の  
等化において合計と等化の順序が与える  
影響について」(発表者 佐藤喜一、柴山直)

## 6. 研究組織

### (1) 研究代表者

柴山 直 (SHIBAYAMA TADASHI)  
東北大学・大学院教育学研究科・教授  
研究者番号：70240752

### (2) 研究分担者

新田 克己 (NITTA KATUMI)  
東京工業大学・大学院総合理工学研究  
科・教授  
研究者番号：60293073

野口 裕之 (NOGUCHI HIROYUKI)  
名古屋大学・大学院教育発達科学研究  
科・教授  
研究者番号：60114815

石井 秀宗 (ISHII HIDETOKI)  
名古屋大学・大学院教育発達科学研究  
科・准教授  
研究者番号：30342934

熊谷 龍一 (KUMAGAI RYUICHI)  
東北大学・大学院教育学研究科・准教授  
研究者番号：60422622

藤本 亮 (FUJIMOTO AKIRA)  
静岡大学・大学院法務研究科・教授  
研究者番号：80300474

藤田 政博 (FUJITA MASAHIRO)  
関西大学・社会学部・准教授  
研究者番号：60377140

佐藤 喜一 (SATO YOSHIKAZU)  
新潟大学・教育・学生支援機構入学セン  
ター・准教授  
研究者番号：00300517

佐藤 誠子 (SATO SEIKO)  
東北大学・大学院教育学研究科・博士研  
究員  
研究者番号：20633655

### (3) 連携研究者

前田 忠彦 (MAEDA TADAHIKO)  
統計数理研究所・データ科学研究系・准  
教授  
研究者番号：10247257

中畝 菜穂子 (NAKAUNE NAOKO)  
国立教育政策研究所・研究員  
研究者番号：00321563

五島 譲司 (GOTO JYOJI)  
新潟大学・教育・学生支援機構 教育支  
援センター・准教授  
研究者番号：90360205