

## 科学研究費助成事業 研究成果報告書

平成 27 年 9 月 25 日現在

機関番号：13901

研究種目：基盤研究(B) (一般)

研究期間：2012～2014

課題番号：24330031

研究課題名(和文) テスト理論による法学テストの能力試験化へ向けての研究

研究課題名(英文) A study on law test as a qualifying test utilizing test theory.

## 研究代表者

藤本 亮 (Fujimoto, Akira)

名古屋大学・法学(政治学)研究科(研究院)・教授

研究者番号：80300474

交付決定額(研究期間全体)：(直接経費) 7,900,000円

研究成果の概要(和文)：TOEFLやTOEICなどで用いられている等化という方法によって、異なった試験の成績を比較することができるようになる。こうした成績測定分野はテスト理論として研究されている。本研究は、テスト理論の見地から、複数の法律学試験において等化を行い、その下でも成績測定が適切に行えるかを検証している。法律学試験は「資格試験」として実施されることが多いが、実質的には一回限りの競争試験となっている。この研究は、本来の意味での資格試験としての法律学試験の可能性を探る基礎研究である。

研究成果の概要(英文)：Equating, used in TOEFL or TOEIC, makes it possible to compare different set of testing. Test theory studies this field of measuring test score. This study, from the perspective of test theory, had equated different sets of law tests to examine if the validity of the equated scores. Law test usually is administered as a “qualifying” test, but in most cases it is indeed a “competitive” one. This study is a basic study for making law test more “qualifying” test as it is supposed to be.

研究分野：法社会学

キーワード：テスト理論 法学テスト 法知識 法学教育 等化

## 1. 研究開始当初の背景

2008年12月の中央教育審議会答申「学士課程教育の構築に向けて」は「学士力」を打ち出したが、その一部として「学習成果の測定・把握」等の調査・研究が言及されている。同年5月には文科省から学術会議に「大学教育の分野別質保証のあり方」についての審議依頼がされ、その後法学を含む「分野別参照基準」の作成が進行中である。今後、法学を学部で学んだ者に必要な知識、能力、マインドについて各大学がその質的保証をすることが求められるようになり、教育成果についての標準化された評価方法が重要となる。OECDの「高等教育における学習成果の評価(AHELO)」の試行研究にも日本は参加している。日本テスト学会は尺度化(標準化)を含むより適正なテスト運用のために近年「テストの開発、実施、利用、管理にかかわる規準」を公表している。このようにテストを含めた教育成果の評価の質を高めることは国内外で大きな流れとなっている。

各種の法学系資格試験では正解率(テスト理論では通過率または困難度)や平均点での検討は行われている。法学の択一式試験は長く司法試験や公務員試験等で用いられており、また、その難易度や正解率についてのみならず、内容的な難問や、知識の有無とは無関係なクイズ解答能力を測定しているような問題は批判もされてきたところである。しかし、こうした検討では、内的一貫性(信頼性)の検証がなされていない。また、受験者集団とそのテスト結果(得点分布や平均点等)に依存するため、異なる機会のテストの間の得点比較は行うことはできない。また、法学の体系性との関係で、限られた問題数によって受験者の法学の学習到達度をいかに計測するのかという妥当性の検討もまだ不十分である。

## 2. 研究の目的

(1) 日本のテスト文化についての先行研究を踏まえつつ、法律学試験の社会的機能についての国際比較も含め、テスト理論、社会学、言語学等からの学際的な検討を行い、その特性を明らかにする。

(2) 高度な知識能力試験としての法律学試験の特性を前提に、法律学試験の択一式で問われるべき内容についての知見を蓄積する。

(3) 法学部生を対象に実験データを収集し、標準化されたスコアが適切に法律学の能力値を測定しているのか多面的に検討する。

## 3. 研究の方法

財団法人商事法務研究会・法学検定試験委員

会より法学検定試験2~4級の実試験データのデータ提供を受け、古典的テスト理論と現代的テスト理論による分析を行う。

法律学を一定程度学習している者を対象に、問題難易度や出題内容について十分に検討した複数の問題セットによる模擬試験実験(同一受験者モデル)を行う。

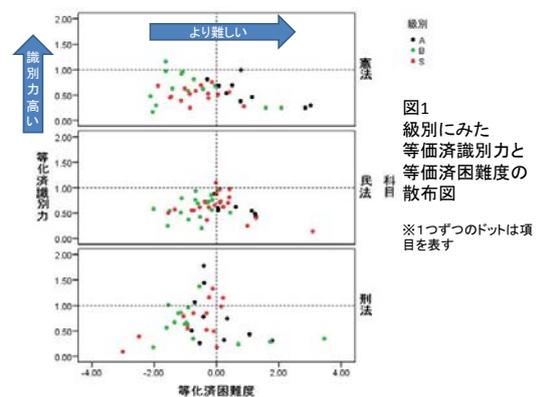
以上により、内容的素点による比較よりも、等化を踏まえたスコアの方がより妥当な能力値推定ができることを明らかにする。

## 4. 研究成果

### A. 法学検定試験の分析

法学検定試験は、ベーシック(旧4級相当)、スタンダード(旧3級相当)、アドバンスト(旧2級相当)の順に難易度が増すように設計されている。ベーシックとスタンダード、スタンダードとアドバンストのそれぞれは重複受験が可能となっている。その重複受験者の解答データを用いた「重複受験者モデル」により、2012年度実試験の憲法・民法・刑法解答データにつき、2パラメータモデルによる分析を行った。まず、項目の特性尺度推定値を計算し、それらに基づいて個人ごとの尺度値推定値を計算した。

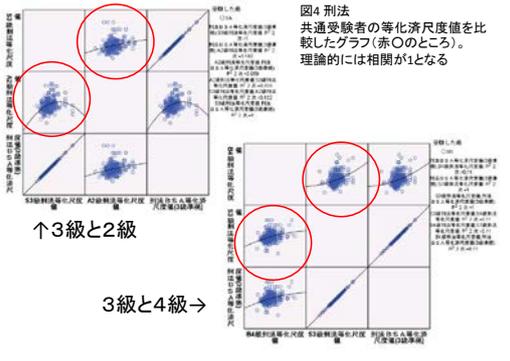
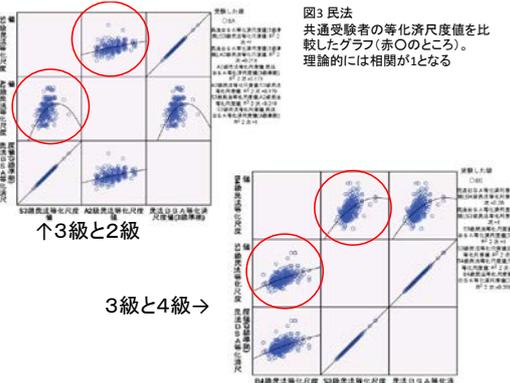
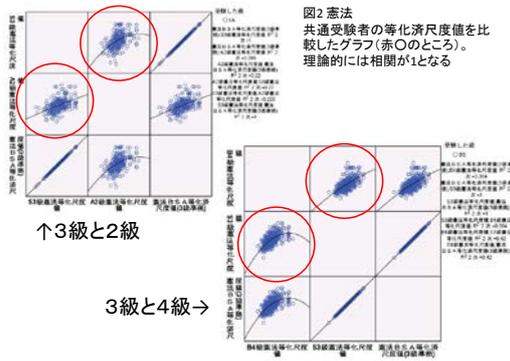
図1は項目ごとの等化済識別力と等化済困難度をプロットしたものである。等化済み尺度スコアは、級ごとに入り組んでいるが、



Basic (4級) Standard (3級) Advanced (2級)の順に難しくなっている弱い傾向がみられる。

これらの特性尺度推定値(識別力と困難度)と個人の解答正誤パターンを用いて、個人の解答パターンが得られる尤度が最大となる $\theta$ の値を個人の特性尺度値の推定値とした。

共通受験者の等化済尺度値を科目別比較したグラフが以下の図2~4である。(赤○かせ)。理論的には相関が1となるが、図に示されたように誤差が含まれる。



憲法については、等価済尺度値間に一定の相関が観察される。民法について、外れ値の影響はあるが全体として弱い相関が観察される。刑法については、とても弱い相関が観察されている。

これらの分析と個別問題の内容的検討から、級別合否ではなく、単一の「法学テストスコア」が得られるようなテスト設計をするためには、以下のような点についての考慮が必要であることが明らかとなった。5肢選択式における「正しいもの」を選択する問題と「誤っているもの」を選択する問題の混在や複雑な組み合わせ問題（肢組み合わせまたは語句組み合わせのパズル式）は、古典的テスト理論の観点からも望ましいものではないが、ここでの分析でも問題項目のパラメータに大きく影響することがわかった。

また、法学テストに特有の問題として、条文知識を尋ねているのか、判例知識を尋ねているのか、さらに判例知識の問題にあっては、ひとつの判例についての5肢なのか、複数の

判例についての5肢なのかという問題パターンの影響もみられる。

現在の法学テストのありかたを前提とするに、安定した法学テスト尺度を構成するためには、まずは古典的テスト理論に基づく丁寧なテスト問題項目設計が重要である。

## B. 実験計画法データの分析

実験計画法においては、これらの分析を踏まえ、問題項目の問題パターンによる影響を排するため、正誤問題に限定して出題することとした。

問題セットの設計にあたっては、法学検定試験の問題の肢を正誤問題に改め、出題内容を体系的に整理し、オリジナル問題が出題された級を参考に作成した。当初、問題セットは、共通項目13項目とセット毎の27項目で憲民刑の3科目について作成し、共通項目モデルでの設計とした。予備的調査において、刑法については、学習程度のばらつきが大きく、精度の高いデータを模擬実験で得ることは困難であるとのことから、憲法・民法の科目について実施することとした。また、ケース数を確保する関係から、共通項目モデルから共通受験者モデルへの変更を行った。

結果として、憲法が54項目、民法が54項目の問題セットを作成した。それぞれの科目においては、27項目ずつセット1とセット2にサブセットに分類されている。それぞれのサブセットは出題分野や難易度（法学検定の出題級と内容的検討による）を調整し、できるだけ平行となるように設計した。

調査協力者はウェブ調査を用いた。ウェブ調査会社の登録者からスクリーニング質問（法学系資格のための勉強をどの程度行っているか）により、本サンプル800、予備サンプル191を選び、最終的に400ケースの解答データを得た。

解答は、単純な正誤ではなく、「正しい」「自信ないが正しいと思う」「自信ないが誤りと思う」「誤り」の選択肢から選ぶように求め、推測によるランダム解答が明らかとなるようにした。

それぞれの項目に記された命題の正誤と合わせて、「確信正解」「推測正解」「推測不正解」「確信不正解」の解答データが得られた。表1はこの回答分布をみたものである。表2は、通常のテストにおける採点と同様に、推測を含む正解も「正解」として分布を見たものである。表3は、「確信正解」のみを正解として分布をみたものである。

表2と表3の欄外に、問題のサブセットごとに信頼性係数を示した。信頼性係数はクロンバッハの $\alpha$ を用い、問題セットが測定している能力が一元的である程度をみる指標である。表2のように通常のテストと同じような採点をした場合の信頼性係数は、憲法でたいへん低くなっている。また、民法のセット1についてもかなり低い値となっている。

表1 問題セットごとの正解分布

サブセット	確信不正解	推測不正解	無答	推測正解	確信正解
憲法セット1	26.4%	23.2%	1.3%	20.3%	28.7%
憲法セット2	28.2%	23.7%	0.9%	21.3%	25.9%
民法セット1	22.1%	19.5%	1.2%	19.9%	37.2%
民法セット2	18.1%	17.4%	4.6%	21.8%	38.2%

表2 問題セットごとの正解分布(推測正解も「正解」)

サブセット	不正解	無答	正解	信頼性係数
憲法セット1	51.0%		49.0%	0.139
憲法セット2	52.9%		47.1%	0.145
民法セット1	42.8%		57.2%	0.360
民法セット2	40.0%		60.0%	0.658

表3 問題セットごとの正解分布(確信正解のみ「正解」)

サブセット	不正解	無答	推測正解	確信正解	信頼性係数
憲法セット1	71.3%			28.7%	0.841
憲法セット2	74.1%			25.9%	0.843
民法セット1	62.8%			37.2%	0.900
民法セット2	61.8%			38.2%	0.920

しかし、「確信正解」のみを正解とした採点方法では、これらのサブセットも高い信頼性係数を示しており、この視点からのテスト性能は低くない。正誤問題の解答にあたっては推測解答が可能であることから、法学の能力をマーク式テストで測定しようとする際にはこの点での考慮も必要であることがわかる。また、このように法の知識が確実に得られているかどうかという点で評価することがテストの性能を高めることがわかる。

サブセット間の正答数得点の関係をグラフにしてみよう。

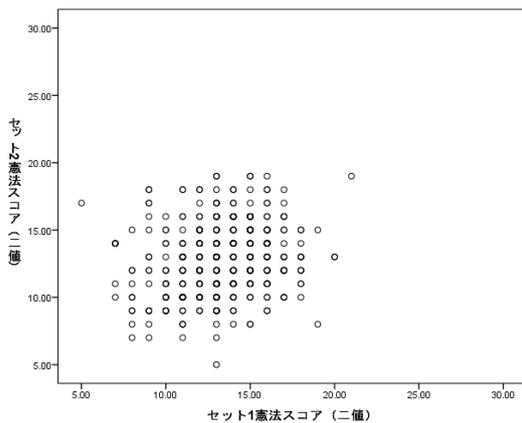


図5 推測正解も「正解」とした場合の正答数

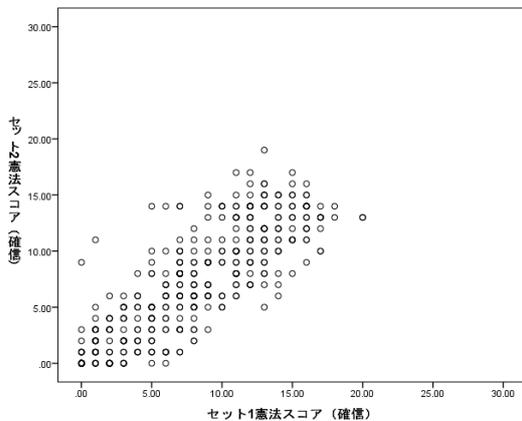


図6 確信正解を「正解」とした場合の正答数

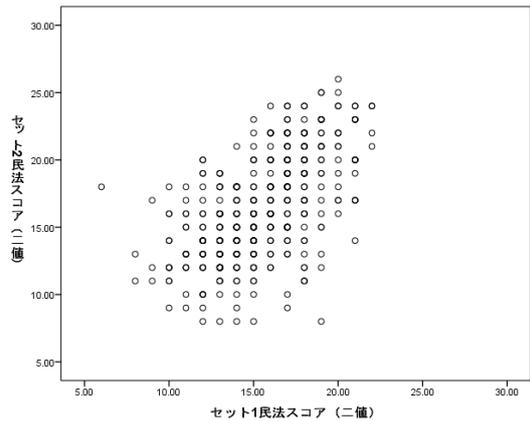


図7

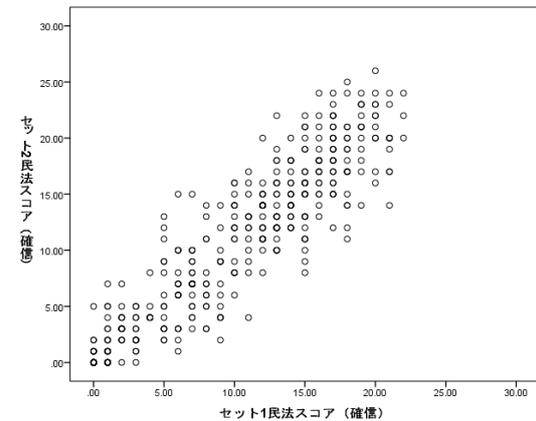


図8

憲法について図5と図6、民法について図7と図8を比較してみると「確信正解」のみを正解として採点した方が異なるテスト間の得点の相関が高くなっていることが一見してわかる。なお、推測正解も正解とした場合と確信正解のみを正解とした場合のそれぞれの正答数スコアのサブセット間の相関係数は表5、表6のとおりである。

表4 採点方法間のスコアの相関(憲法)

		セット1憲法スコア(二値)	セット2憲法スコア(二値)	セット1憲法スコア(確信)	セット2憲法スコア(確信)
セット1憲法スコア(二値)	Pearson Correlation	1	.210**	.410**	.231**
	Sig. (2-tailed)		.000	.000	.000
	N	400	400	400	400
セット2憲法スコア(二値)	Pearson Correlation	.210**	1	.110	.254**
	Sig. (2-tailed)	.000		.028	.000
	N	400	400	400	400
セット1憲法スコア(確信)	Pearson Correlation	.410**	.110	1	.841**
	Sig. (2-tailed)	.000	.028		.000
	N	400	400	400	400
セット2憲法スコア(確信)	Pearson Correlation	.231**	.254**	.841**	1
	Sig. (2-tailed)	.000	.000	.000	
	N	400	400	400	400

\*\* . Correlation is significant at the 0.01 level (2-tailed).

\* . Correlation is significant at the 0.05 level (2-tailed).

表5 採点方法間のスコアの相関(民法)

		セット1民法スコア(二値)	セット2民法スコア(二値)	セット1民法スコア(確信)	セット2民法スコア(確信)
セット1民法スコア(二値)	Pearson Correlation	1	.583**	.628**	.560**
	Sig. (2-tailed)		.000	.000	.000
	N	400	400	400	400
セット2民法スコア(二値)	Pearson Correlation	.583**	1	.540**	.680**
	Sig. (2-tailed)	.000		.000	.000
	N	400	400	400	400
セット1民法スコア(確信)	Pearson Correlation	.628**	.540**	1	.915**
	Sig. (2-tailed)	.000	.000		.000
	N	400	400	400	400
セット2民法スコア(確信)	Pearson Correlation	.560**	.680**	.915**	1
	Sig. (2-tailed)	.000	.000	.000	
	N	400	400	400	400

\*\* . Correlation is significant at the 0.01 level (2-tailed).

この分析からもやはり古典的テスト理論に基づいたテスト設計の重要性もまた明らかになったと考えられるのである。

#### C. 研究成果を受けての今後の研究の展開

本研究で得られた成果については、実験計画法データのIRT分析による検討を含め、現在図書としての出版を準備中である。

また、多くの法学系資格試験の問題が多肢選択式となっていることから、多肢選択式において、推測解答 (Guessing) を排除するテストのあり方について、法律学知識の体系性をふまえながら、今後も研究を継続していく予定である。

#### 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 件)

熊谷龍一・野口裕之 (2012) 「推定母集団分布を利用した共通受験者法による等化係数の推定」日本テスト学会誌 8(1)、9-18 頁。

藤本亮 (2013) 「法学分野の参照基準と法学教育」静岡法務雑誌 6号、41-65 頁。

[学会発表] (計 件)

[図書] (計 件)

出版準備中

[産業財産権]

なし

[その他]

ホームページ等

#### 6. 研究組織

##### (1) 研究代表者

藤本亮 (FUJIMOTO, Akira)

名古屋大学・法学研究科・教授

研究者番号：80300474

##### (2) 研究分担者

野口裕之 (NOGUCHI, Hiroyuki)

名古屋大学・教育発達科学研究科・教授

研究者番号：60114815

藤田政博 (FUJITA, Masahiro)

関西大学・社会学部・准教授

研究者番号：60377140

堀田秀吾 (HOTTA, Shugo)

明治大学・法学部・教授

研究者番号：70330008

小谷順子 (KOTANI, Junko)

静岡大学・人文社会科学部・教授

研究者番号：40359972

宮下修一 (MIYASHITA, Shuichi)

静岡大学・法務研究科・教授

研究者番号：80377712

吉川真理 (YOSHIKAWA, Mari)

静岡大学・法務研究科・教授

研究者番号：70289750

正木祐史 (MASAKI, Yushi)

静岡大学・法務研究科・教授

研究者番号：70339597