

【原著】

受験者属性別の教科科目フリー型総合試験成績の比較と特徴

伊藤圭，荒井清佳，桜井裕仁（大学入試センター研究開発部），杉澤武俊（新潟大学）

教科科目フリー型総合試験は幅広い層の受験者を対象とするため、受験者属性間(学部系統，性別)における問題の公平性の確保が一つの課題である。本研究では，モニター調査のデータに基づき，属性間の一般学力差を考慮した上で属性が総合試験成績に及ぼす効果を調べた。その結果，情報把握力や論理的思考力を測る試験は理系および男性にやや有利な傾向が見られ，読解力や表現力を測る試験は文系にやや有利な傾向が見られた。

1 教科科目フリー型総合試験と受験者属性

教科科目フリー型総合試験は，特定の教科科目の知識に依存しないという特徴から，様々な社会的，教育的背景を有した幅広い受験者層に適した試験であることが期待される（藤井ほか，2002；林・伊藤，2010）。例えば，2003年から2010年まで大学入試センターが行ってきた法科大学院適性試験は，法学既修者だけではなく幅広い学部・学科からの受験者を想定しており，試験の内容や形式も典型的な教科科目フリー型であると考えられる。また，大学の医学部学士編入学試験においても，医学知識を必要としない教科科目フリー型の能力試験や適性試験を採用することがある。これらの試験の作題上の課題の一つは，特定の属性を有した受験者にだけ有利または不利に働くような問題をできるだけ排除することである。そこで，本稿では，過去に行われた総合試験に関するモニター調査のデータに基づき，受験者属性がどの程度総合試験成績に影響を及ぼすかについて基礎的な解析を試みた。

2 モニター調査の概要

2.1 試験の構成

大学入試センター研究開発部では，教科科目フリー型総合試験の特徴を調べるために，

情報把握力および論理的思考力を測ることを意図した問題（第1領域と呼ぶ）とコミュニケーション力，読解力，表現力を測ることを意図した問題（第2領域と呼ぶ）の2種類の総合試験を作成し，これまでに計3回のモニター調査を実施してきた（大学入試センター研究開発部，2006，2010，2011）。但し，その後の分析で，第2領域ではコミュニケーション能力が十分に測定できていないことが判明している（伊藤ほか，2006）。

平成17年には，全国の国公立大学47校の学部3～4年生753名に対して第1領域（35項目，50点），第2領域（35項目，50点）の試験を実施した。さらに，受験者の一般学力を代表するものとして，過去の大学入試センター試験の問題から比較的識別力の高い適当な問題を選び，再編集した英語の試験も実施した。以下では，この調査をH17調査と呼ぶ。

平成18年には，東京都内の国立大学5校の学部1年生348名に対して，問題数をH17調査に比べて約半分にした短縮版の試験〔第1領域（15項目，25点），第2領域（17項目，25点）〕を実施した。この調査では，総合試験の他に，受験者の一般学力を代表するものとして，平成18年度大学入試センター試験の主要8教科科目の試験も実施した。以下では，これをH18調査と呼ぶ。

平成19年には、東京都内の国立大学5校の学部1年生321名に対してH18調査と同じ短縮版の試験を実施した。この調査においても、一般的な学力を代表するものとして、平成19年度大学入試センター試験の主要8教科科目の試験を実施した。以下では、これをH19調査と呼ぶ。

2.2 受験者属性

本研究で取り上げる受験者属性は学部系統（文系と理系の2種）と性別（男性と女性の2種）とする。

H17調査では、全受験者753名の中から、受験者の所属学部・学科から判断して、文系225名、理系293名を抽出した。残りの235名はほとんどが医学部等に所属する学生のため、文系か理系かを明確に区別できないものと考え、分析から除いた。文系と理系を合わせた518名の性別内訳は男性312名、女性206名である。

次に、H18調査では、受験者の所属学部・学科および大学入試受験時の大学入試センター試験受験科目から判断して、文系158名、理系190名に分けた。性別内訳は男性289名、女性59名である。

最後にH19調査も同様にして文系148名、理系173名に分けた。性別内訳は男性260名、女性61名である。

H17調査では全受験者の三分の一弱が医学部等に所属しているが、これは総合試験を医学部学士編入学試験として用いる場合の特性を調べるため、学部3～4年生の医学系の受験者を相当数募集したことによる。H18、H19調査の医学系の受験者は、調査時点では教養課程の段階にあり、医学系の学部における専門課程を履修している者はいない。

3 受験者属性別の成績比較について

3.1 総合試験成績と一般学力

属性別に分けられた受験者集団の総合試験

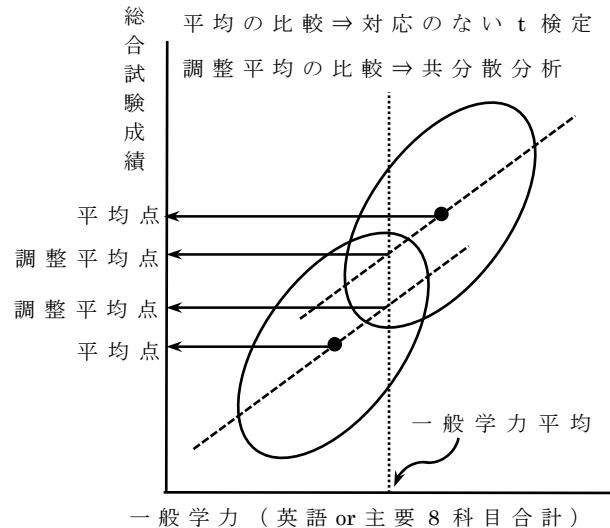


図1 平均点の比較と調整平均点の比較

成績を比較する方法として、試験の平均点を比較することが考えられる。しかしながら属性別に分けられた集団の間に平均点差が存在したとしても、それが属性による有利不利に起因したものなのか、それとも平均点に影響を及ぼす他の要因によるものなのかが自明ではない場合も考えられる。そのような場合には、比較する集団を注目している属性以外の要因に関してできるだけ均質になるように条件づけた上で、集団の等質性や能力の同等性を仮定して比較するなどの工夫が必要となる。

本研究で取り扱う総合試験は能力試験の一つであり、教科科目学力と一定の関係を有していると考えられる（伊藤ほか，2010a）。実際に、総合試験に関するいくつかの先行研究では、総合試験と教科科目試験の相関係数が0.2台半ばから0.5台前半程度になることが確認されている（伊藤，2006a；伊藤ほか，2006）。したがって、一般的な学力は総合試験成績へある程度の（少なくとも受験者属性と比較し得る程度の）影響を及ぼす要因の候補の一つであると考えられる。そこで、本研究では一般学力の影響を考慮するために、H17調査における英語およびH18、H19調査における大学入試センター試験主要8教科科目の総合点を総合試験の平均点に関する共変

量として取り上げ、共分散分析的手法により受験者属性の総合試験成績への影響を調べることにした。

実際には、まず共変量を考慮しない場合の比較、すなわち総合試験の通常平均点の属性間比較を行い、その後、共変量を考慮して算出した調整平均点の属性間比較を行った。両者の違いを概念図として図1に示す。図は一般学力と総合試験得点の散布図であり、図中の二つの楕円はそれぞれ文系と理別、または男性と女性の受験者集団を便宜的に表したもので、両者を合わせたものが全受験者を表す。右上がりの点線は各属性の回帰直線であり、垂直の点線は一般学力を表す共変量の平均値の位置を表す。全受験者の一般学力を平均値に条件付ける場合、比較する二つの集団の平均点、すなわち図中の黒点を回帰直線に沿って一般学力平均まで平行移動することになる。これにより、全受験者が等しい一般学力であると仮定した場合の総合試験の平均点（調整平均点）を算出することができる。この調整平均点を比較することにより、一般学力の影響を除いた状態で、受験者属性が総合試験成績に及ぼす影響を調べることができる。

具体的には、平均点の比較および調整平均点の比較は以下のように行った。

3.2 共変量を考慮しない場合の2群の比較

まず、H17～H19調査のそれぞれについて、第1領域得点および第2領域得点の文系と理系の平均点差および男性と女性の平均点差の有意性を2群の等分散性を考慮した対応のないt検定により調べた。ただし、H18調査における第1領域の得点については、男女間で有意な分散の差が認められたため（ $F=7.345$ ， $p=0.007(<0.01)$ ），等分散性を仮定しないWelchの検定を行った。

3.3 共変量の設定

次に、比較対象の属性間にそもそも一般的

な学力差がある場合は、総合試験の平均点に関する共変量として一般学力を次のように仮定した。H17調査では、実施上の制約から多数の学科試験を行うことができなかったため、主要な教科科目と比較的高い相関があり、各科目の中間的性質を有している英語の試験を実施し、その得点を一般学力を表す尺度として採用した（池田，1983；伊藤・荒井，2010，2011）。

H18およびH19調査では、大学入試センター試験主要8教科科目の得点に対して分散共分散行列に基づく主成分分析を行い、第1主成分得点を算出した。第1主成分は、多くの場合において、全ての変量を総合した量的な性質を持つことが知られている（奥野ほか，1981；田中・脇本，1983；伊藤，2006b）。実際に、大学入試センター試験の主要教科科目の主成分分析においても第1主成分得点が総合点としての性質を有していることが確認されている（池田，1983；伊藤・荒井，2010，2011）。このことから、第1主成分得点を一般学力を表す尺度として採用した。

一般に、科目 i （ $i=1,2,\dots,8$ ）の標準化されていない得点（素点）を x_i ，得点 x_i に対する重み係数を w_i とすると、重み係数の平方和（ $\sum_{i=1}^8 w_i^2$ ）が一定という条件の下で、主成分得点は $f = \sum_{i=1}^8 w_i x_i$ と表すことができる。重み係数の平方和の値は任意の定数に定めることができるので、今回の分析では8

（ $\sum_{i=1}^8 w_i^2 = 8$ ）とした。これは、教科科目の単純な合計点（800点満点）の場合、各教科科目の重み係数が1で、平方和が8であることに便宜的に合わせたものである。主成分得点と合計点において重み係数の平方和を等しく置いたとしても、厳密に両者を比較可能にはできないが、実際のデータでは両者の得点は同じオーダーとなり、また相関係数も十分に大きいため、大まかではあるが主成分得点から単純な合計点の大きさの程度を見当づけることができる（伊藤，2006b）。

表 1 第 1 領域得点の群間（学部系統間，男女間）の比較結果

第1領域（文理，調整前）						
	H17調査		H18調査		H19調査	
	平均	標準偏差	平均	標準偏差	平均	標準偏差
文系	23.89	7.97	14.46	4.66	13.41	5.04
理系	26.84	8.45	14.97	4.72	16.20	4.69
<i>t</i> 値	<i>t</i> =4.04 (<i>p</i> <0.01)		<i>t</i> =1.01 (<i>p</i> =0.314)		<i>t</i> =5.12 (<i>p</i> <0.01)	

第1領域（文理，調整後）						
	H17調査		H18調査		H19調査	
	調整平均	標準偏差	調整平均	標準偏差	調整平均	標準偏差
文系	22.85	6.85	14.81	4.13	14.10	3.96
理系	27.66	6.70	14.56	3.74	15.64	3.89
<i>F</i> 値	<i>F</i> =55.76 (<i>p</i> <0.01)		<i>F</i> =0.37 (<i>p</i> =0.544)		<i>F</i> =11.58 (<i>p</i> <0.01)	

第1領域（男女，調整前）						
	H17調査		H18調査*		H19調査	
	平均	標準偏差	平均	標準偏差	平均	標準偏差
男性	27.32	8.01	15.03	4.50	15.78	4.71
女性	22.89	8.20	13.32	5.38	11.18	4.77
<i>t</i> 値	<i>t</i> =6.09 (<i>p</i> <0.01)		<i>t</i> =2.27 (<i>p</i> =0.026)		<i>t</i> =6.84 (<i>p</i> <0.01)	

第1領域（男女，調整後）						
	H17調査		H18調査		H19調査	
	調整平均	標準偏差	調整平均	標準偏差	調整平均	標準偏差
男性	27.16	7.10	14.75	3.84	15.21	3.97
女性	23.18	6.90	14.93	4.42	13.63	3.90
<i>F</i> 値	<i>F</i> =39.79 (<i>p</i> <0.01)		<i>F</i> =0.00 (<i>p</i> =0.978)		<i>F</i> =6.62 (<i>p</i> =0.011)	

*等分散性を仮定しないWelchの検定による。

3.4 共変量を考慮した場合の2群の比較

共変量を設定した後，H17～H19調査のそれぞれについて，第1領域得点および第2領域得点の学部系統間（文理間），男女間における差について共分散分析を行った。分析は三つの調査（H17，H18，H19），二つの総合試験領域（第1領域，第2領域），二つの属性（学部系統，性別）からなる12通りの組み合わせのそれぞれの場合について行った。すべての組み合わせについて総合試験得点の共変量への回帰係数を調べたところ，学部系統間，男女間ともに有意水準1%で有意な差は認められなかった。そこで，総合試験得点

から共変量への回帰による変動分を差し引いた調整得点 [調整前の属性別素点を y ，それに対する共変量を x ， x の全体の平均を μ ，回帰係数を β とすると， $y + \beta(x - \mu)$ で表される。] を算出し，調整平均点の比較を行った。すなわち，各受験者の共変量の値を平均に条件づけた場合の総合試験得点を比較することにより，同等の一般学力の下で2群を比較しているものとみなすことにした。

4 結果

第1領域得点の群間（学部系統間，男女間）の比較結果を表1に，同様に第2領域得点の

表 2 第 2 領域得点の群間（学部系統間，男女間）の比較結果

第2領域（文理，調整前）						
	H17調査		H18調査		H19調査	
	平均	標準偏差	平均	標準偏差	平均	標準偏差
文系	31.11	5.75	13.61	3.94	12.86	4.21
理系	29.95	6.02	12.72	4.25	12.42	3.85
<i>t</i> 値	<i>t</i> =2.22 (<i>p</i> =0.027)		<i>t</i> =2.00 (<i>p</i> =0.047)		<i>t</i> =0.97 (<i>p</i> =0.335)	

第2領域（文理，調整後）						
	H17調査		H18調査		H19調査	
	調整平均	標準偏差	調整平均	標準偏差	調整平均	標準偏差
文系	30.55	5.14	13.82	3.71	13.28	3.77
理系	30.47	5.19	12.42	3.77	12.06	3.45
<i>F</i> 値	<i>F</i> =0.02 (<i>p</i> =0.901)		<i>F</i> =11.51 (<i>p</i> <0.01)		<i>F</i> =8.64 (<i>p</i> <0.01)	

第2領域（男女，調整前）						
	H17調査		H18調査		H19調査	
	平均	標準偏差	平均	標準偏差	平均	標準偏差
男性	30.30	6.00	13.13	4.12	12.82	4.01
女性	30.67	5.81	13.07	4.21	11.80	3.99
<i>t</i> 値	<i>t</i> =7.00 (<i>p</i> =0.484)		<i>t</i> =0.11 (<i>p</i> =0.910)		<i>t</i> =1.77 (<i>p</i> =0.078)	

第2領域（男女，調整後）						
	H17調査		H18調査		H19調査	
	調整平均	標準偏差	調整平均	標準偏差	調整平均	標準偏差
男性	30.18	5.34	12.95	3.82	12.44	3.64
女性	30.88	4.87	14.11	3.71	13.36	3.58
<i>F</i> 値	<i>F</i> =2.18 (<i>p</i> =0.141)		<i>F</i> =3.61 (<i>p</i> =0.058)		<i>F</i> =2.80 (<i>p</i> =0.095)	

群間の比較結果を表 2 に示す。表 1，表 2 ともに，上から 1，3 段目の表が調整前の素点の平均とその差に関する *t* 検定の結果であり，上から 2，4 段目の表が共変量（一般学力）を考慮した場合の調整得点の平均とその共分散分析の結果である。ただし，H17 調査の試験は 50 点満点，H18，H19 調査の試験は 25 点満点である。なお，有意水準 5% で有意な平均点差が認められた部分の数値を太字で示し，さらに有意水準 1% でも有意な差が認められた部分を網掛けで示してある。

4.1 第 1 領域の結果

第 1 領域における文系と理系の平均点差を

調べた結果，H17 および H19 調査において，調整前，後ともに有意水準 1% で有意な差が見られた。いずれも理系の方が平均点が高い。調整後の有意な平均点差は 1.5 ～ 4.8 であり，およそ小問 2 ～ 4 題の得点に相当する程度であった。

次に，第 1 領域における男性と女性の平均点差を調べた結果，H17 調査の調整前，後，H19 調査の調整前において有意水準 1% で有意な差が見られた。さらに，有意水準 5% まで許容すると H18 調査の調整前，H19 調査の調整後においても有意な差となった。いずれも男性の方が平均点が高い。調整後の有意な平均点差は 1.6 ～ 4.0 であり，これも小問 2

～4題ほどの得点に相当する程度であった。

4.2 第2領域の結果

第2領域における文系と理系の平均点差を調べた結果、H18およびH19調査の調整後において有意水準1%で有意な差が見られた。さらに有意水準5%まで許容するとH17およびH18調査の調整前においても有意な差となった。いずれも文系の方が平均点が高い。調整後の有意な平均点差は1.2～1.4であり、これはおよそ小問1題の得点に相当する程度であった。

次に、第2領域における男性と女性の平均点差を調べた結果、H17～H19のいずれの調査においても、有意水準5%で有意な差は見られなかった。平均点差は大きくても1.2程度であり、これも小問1題程度に相当する得点差に収まっている。

5 まとめと今後の課題

教科科目フリー型総合試験は幅広い受験者層を対象とすることから、受験者属性間における試験問題の公平性が求められる。そこで本研究では、先に述べた第1領域試験と第2領域試験について、計3回のモニター調査のデータを用いて、学部系統間（文理）と男女間での平均点の比較を行い、受験者属性の試験成績への影響を調べた。特に、本研究では総合試験得点と相関する一般学力の効果を考慮するため、英語または複数の学科試験の総合点を共変量とした共分散分析を行った。

分析の結果、第1領域については、全体として理系および男性に有利な傾向が見られた。第2領域については文系に有利な傾向が見られたが、男女間には差が見られなかった。結果的に、情報分析や論理的な結論の導出等は理系および男性が得意であり、意味、文脈の把握、表現の読みかえ、相対的価値判断等は文系が得意であるといった、ステレオタイプな描像を想起させるものとなった。

調整後の得点において見られた有意な平均点差は、第1領域の学部系統間および男女間で小問2～4題の得点に相当し、第2領域の学部系統間で小問1題の得点に相当する程度であった。総合試験の満点が短縮版冊子で25点、通常版冊子で50点であり、測定レンジがやや狭いことから、第1領域で見られた平均点差は満点に対して5～10%を占める比較的大きな割合となった。既に、項目レベルの分析を行った先行研究において、受験者属性間に困難度差が認められた項目が報告されているので（伊藤ほか、2010b）、これらの項目を除去した場合の冊子レベルでの受験者属性間の平均点差についても調べる必要があるであろう。

本研究では総合試験成績に関する共変量として教科科目の総合点を用いて一般学力を仮定したが、伊藤ほか（2010a）の研究では教科科目学力と総合試験成績の違いを表す因子と問題解決や課題遂行に必要な基礎的能力との関連性が確認されている。これらの基礎的な能力の特徴に基づいて受験者を群分けし、群間での総合試験成績を比較することにより、学部系統、性別以外の影響を調べることが可能と思われる。

本稿では試験の平均点差のみを分析したが、得点分布に着目した分析も必要と考えられる。総合試験得点および共変量について天井効果、床効果、多変量はずれ値の効果等をチェックするとともに、分布の差の分析を行うことにより、低得点域、高得点域といった得点領域ごとの特徴を明らかにすることができると考えられる。また、学部系統と性別を同時に取り扱った2要因の分析についても今後の課題としたい。

問題作成の過程で属性間の差異をできるだけ無くすように努めたとしても、実際には問題冊子単位である程度の差異が生じてしまうことは少なくないと思われる。試験の作成や実施における実際の制約下でどの程度の差異

を許容するかは、その試験の用途や社会的影響の大きさに関係するだけでなく、属性間の差異の意味をどのように解釈するかということにも関係すると考えられる。例えば教科科目フリー型総合試験の場合、「属性間の差異が教科科目や出身分野に固有の知識の有無による差でなければ良いとする」、「属性間の差異が生得的または社会的に形成された得意・不得意による差であるならば、試験問題の善し悪しとは無関係とする」、「属性間に有利・不利があっても、入学後の学習活動に本質的に必要な能力であるならば、調整されるべき差異ではないとする」といった問題が挙げられるであろう。

参考文献

- 大学入試センター研究開発部 (2006). 『平成 15-17 年度共同研究報告書総合試験の分析的研究』 .
- 大学入試センター研究開発部 (2010). 『平成 18-20 年度プロジェクト研究報告書新しい枠組みとしての総合試験の実証的研究』 .
- 大学入試センター研究開発部 (2011). 『平成 21-22 年度プロジェクト研究報告書新しい枠組みとしての総合試験の実証的研究 (2)』 .
- 藤井光昭・柳井晴夫・荒井克弘一編著 (2002). 『大学入試における総合試験の国際比較－我が国の入試改善に向けて』 多賀出版 .
- 林篤裕・伊藤圭 (2010). 「総合試験の実態調査」 『大学入試研究ジャーナル』 20, 57-61.
- 池田央 (1983). 「共通 1 次試験の教科・科目間の相関」 『大学入試フォーラム』 2, 54-61.
- 伊藤圭 (2006a). 「総合試験と教科・科目別試験との相関についての分析」 『平成 18 年度大学入試センター試験モニター調査研究報告』 大学入試センター研究開発部, 89-103.
- 伊藤圭 (2006b). 「主成分分析」 『SPSS による統計データ解析－医学・看護学, 生物学, 心理学の例題による統計学入門－』 現代数学社, 第 7 章, 173-201.
- 伊藤圭・荒井清佳 (2010). 「教科科目の相関についての分析」 『平成 22 年度大学入試センター試験モニター調査研究報告』 大学入試センター研究開発部, 141-150.
- 伊藤圭・荒井清佳 (2011). 「平成 23 年度大学入試センター試験モニター調査教科科目相関分析」 『平成 23 年度大学入試センター試験モニター調査研究報告』 大学入試センター研究開発部, 201-223.
- 伊藤圭・林篤裕・椎名久美子・大澤公一・石井秀宗・柳井晴夫・田栗正章・岩坪秀一・赤根敦・麻生武志・岩堀淳一郎・内田千代子・川崎勝・齋藤宣彦・武田龍司 (2006). 「医学部学士編入学者選抜のための総合試験の開発とその評価」 『大学入試センター研究紀要』 35, 49-108.
- 伊藤圭・林篤裕・椎名久美子・田栗正章・小牧研一郎・柳井晴夫 (2010a). 「学科試験および科目得意度との比較による総合試験の妥当性の検証」 『日本テスト学会誌』 6, 114-123.
- 伊藤圭・大久保智哉・柳井晴夫 (2010b). 「項目困難度による総合試験の問題内容分析」 『大学入試研究ジャーナル』 20, 63-73.
- 奥野忠一・久米均・芳賀敏郎・吉澤正 (1981). 『多変量解析法《改訂版》』 日科技連出版社 .
- 田中豊・脇本和昌 (1983). 『多変量統計解析法』 現代数学社 .