

「言語運用力」試験の開発と検討

—大学新入生に対する 2013 年度調査の結果とその分析—

伊藤圭，荒井清佳，椎名久美子，宮埜寿夫，桜井裕仁，小牧研一郎（大学入試センター），
田栗正章（中央大学），安野史子（国立教育政策研究所）

近年増加した AO 入試や推薦入試による大学入学志願者の基礎的な学力を，教科科目の学科試験とは異なる観点から評価することを目的として，実践的な言語運用力を測る問題を試作し，モニター調査による評価を行った。その結果，想定する受験者に対して，試作問題が概ね適した困難度と識別力を有することが確認された。また，問題項目の解答データに対して主成分分析を行った結果，試験の構成概念（試験で測定しようとしている能力）に対応する因子的構造が確認された。

1 はじめに

近年，大学進学率が上昇し，大学入学志願者の社会的，教育的背景が多様化した。AO 入試や推薦入試による大学入学者数は教科科目の学科試験を主体とした，いわゆる一般入試による入学者数と同程度の規模にまで増加している（文部科学省，2013）。AO 入試や推薦入試を経た大学入学者は，大学入試センター試験や個別大学の学科試験などの教科科目型試験を受験していない場合が多く，これらの学生の学力を適切に評価する方法が求められている。

このような状況に鑑み，大学入試センター研究開発部では，2011 年度から，実践的な言語運用力や数理的な理解力，思考力，問題解決力等を評価する試験問題の開発に関する調査研究を進めている。2012 年度には，試作した「言語運用力」と「数理分析力」の問題のモニター調査を行い，より良質な問題の開発に向けた予備的検討を行った（椎名ほか，2014；桜井ほか，2014a；伊藤ほか，2014）。

2013 年度には，2012 年度の調査の分析結果に基づいて問題を精選，改良し，再構成した「言語運用力」と「数理分析力」の試験

について，AO 入試や推薦入試を経た入学者が多い大学・短期大学の 1 年生を対象としたモニター調査を行った。本稿では 2013 年度に行われた調査のうち，「言語運用力」試験に関する調査の概要と分析結果を報告する。

2 「言語運用力」試験の構成

「言語運用力」試験は，個別の教科科目の学習到達度を測る従来の学科試験とは異なる観点から，実践的な言語運用能力を測ることを基本的な目的としている。具体的には，大学入学後の講義で教科書や参考文献を読んで理解するのに必要となる読解力を評価するために，表 1 に示した L1 ～ L3 の能力を測定することを意図して問題を作成した。

例えば，問題文中に提示されている個々の

表 1 言語運用力の能力分類

ラベル	測定する能力
L1	情報の把握:細かい情報も含め，文章内の情報を正しく読み取る能力
L2	内容の理解:文章の内容の理解や解釈を行う能力
L3	推論と推察:内容の理解にとどまらず，推測，評価，判断等を行う能力

表 2 「言語運用力」の冊子構成と問題内容

大問番号	問題の内容	素材文	設問	解答番号	ラベル
第1問	紛らわしい表現の理解	日本語	問1	1	L2
			問2	2	L2
第2問	会話の内容の正確な読み取り	英語	問1	3	L1, L2
			問2	4	L1
第3問	会話の文脈の理解に基づく適切な応答	英語	問1	5	L2, L3
			問2	6	L2, L3
第4問	会話文から読み取った情報を 地図に適用して理解する	日本語	問1	7	L1
			問2	8	L1
			問3	9	L2
第5問	正しい推論を選ぶ	日本語		10	L3
第6問	会話の内容からの状況の推測	英語	問1	11	L3
			問2	12	L2
第7問	長文の読み取りとそれに基づく類推	日本語	問1	13, 14	L1
			問2	15	L1
			問3	16	L3
第8問	文章の内容の正確な読み取り	英語		17	L2

具体的事実（乗り物の発着時刻や対話が行われている場所など文脈によって意味が変化しない事項）などを正しく読み取る能力は L1 に該当する。また、問題文の文脈を把握し、その内容を別の表現に置き換えたり、図表を対応づけたりする能力などは L2 に該当する。さらに、問題文から得られた情報と既に身につけている知識との関連づけ、条件や程度の評価、複数の可能性の比較などを通して、相対的により適切な判断を行う能力などは L3 に該当する。但し、これらはいくまで一例であり、表 1 に示す能力が上記の例に限定されるものではない。

表 2 に問題冊子の構成を示す。素材文が日本語で提示された問題（以下、日本語問題）が 4 問、英語で提示された問題（以下、英語問題）が 4 問の合計 8 問の大問から成る。解答時間は 40 分、解答形式は大学入試センター試験と同様のマークシートによる多枝選択式である。設問数は 16 であるが、第 7 問の間 1 が選択枝から二つの解答を選ぶ問題であるため、採点の最小単位となる採点項目数は 17 である。

出題した問題は、書かれていることを正確に読み取る力や、単語や文章の意味を文脈から類推する力、文脈にしたがって図表を読み

取る力、文章の論理構造を読み取る力を測定しようとするものなど、表 1 の能力分類の偏りがないように考慮して組み合わせてある。また、実際の日常場面も題材として採用し、会話文の問題も含まれている。英語問題は過去に行われた大学入試センター試験の「英語リスニング」の問題を利用し、音声で提示された問題を文字による提示に置き換えたり、選択枝を一部変更したりしたものである。

3 モニター調査の概要

2013 年度に行われたモニター調査では、いずれも AO 入試や推薦入試を経た入学者が多いひとつの短期大学と四つの大学、合わせて 5 大学（それぞれ U_1 , U_2 , U_3 , U_4 , U_5 と表記する）の入学直後の 1 年生を対象とした。調査では「言語運用力」試験の他に「数理分析力」試験と基本的な能力資質に関するアンケートも行ったが、それらの分析については別途報告する予定である。受験者数の内訳は表 3 に示すとおりである。

4 分析結果

試験の採点は、採点単位である採点項目に対して正答の場合は 1 点、不正答の場合は 0 点を与える 2 値採点（0/1 採点）とし、本稿

表3 基本統計量の計算結果

	全体	U ₁ 短大	U ₂ 大学	U ₃ 大学	U ₄ 大学	U ₅ 大学
受験者数	2208	145	152	1059	518	334
満点(採点項目数)	17	17	17	17	17	17
平均得点	7.96	8.00	7.49	7.59	8.18	8.99
平均得点率	0.47	0.47	0.44	0.45	0.48	0.53
中央値	8	8	7	7	8	9
第1四分位数	6	6	5	5	6	7
第3四分位数	10	10	10	10	10	12
標準偏差	3.25	3.03	3.29	3.19	3.26	3.22
最大値	17	15	15	17	17	16
最小値	0	0	2	0	1	0
α 係数	0.68	0.65	0.70	0.67	0.68	0.68

ではこのデータを用いて分析を行う。以下、問題冊子全体の得点（採点項目の得点を合計した得点）について言及するときは、単に“得点”と記す。

4.1 基本統計量の計算結果と得点分布

表3に各大学の受験者および全ての大学を合わせた全受験者についての基本統計量の計算結果を示す。また、図1は全受験者の得点の度数分布、図2は全受験者の得点の累積相対度数分布、図3は全受験者と各大学の受験者の得点分布を箱ひげ図で表したものである。

平均得点率は低い大学で4割前半、高い大学で5割前半であるが、全体としては5割弱程度である。得点分布を見ると、やや低得点側への偏りが見られるが、易しすぎることも、難しすぎることもない試験であったと考えられる。

α 係数は受験者全体および各大学ともに0.6台後半～0.7の範囲にある。項目数が17と少なめであることを考慮すると、比較的高い信頼性であると思われる。

「言語運用力」試験と同時に行われた「数理分析力」試験（採点項目数25）との比較のために、今回の「言語運用力」試験で用いた17項目と平均的な性質が同等とみなせる

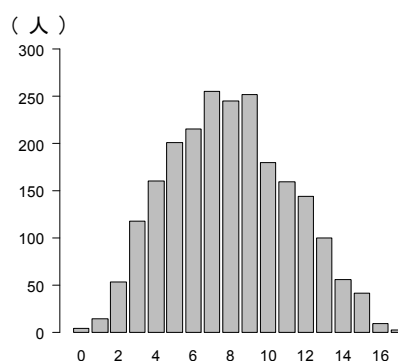


図1 全受験者の得点分布

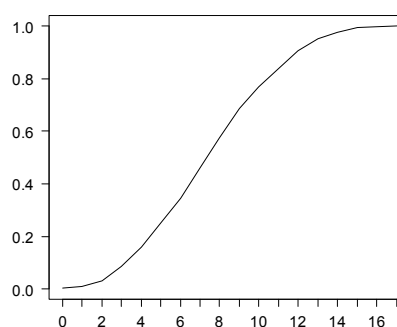


図2 全受験者の得点の累積相対度数分布

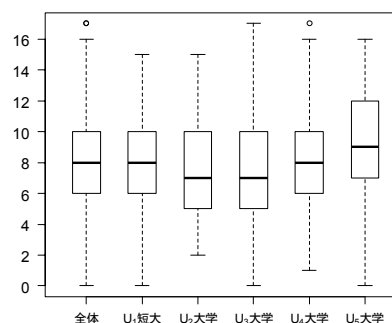


図3 各大学の得点分布

項目（項目間相関係数の平均が同じになる項目）を8項目追加して「数理分析力」試験と同じ採点項目数の冊子を組んだと仮定し、

Spearman-Brown の公式により補正した α 係数を求めると、受験者全体では0.76、大学別では0.73~0.77の範囲の値になった。「数理分析力」試験の α 係数は受験者全体で0.89、大学別では全て0.88であり（桜井ほか，2015），「言語運用力」試験の方がやや低い値となったが概ね許容できる程度の信頼性が得られたと考えられる。

4.2 各設問の正答率と識別力

表4に全受験者の各設問に対する正答率と識別力を示す。表中の解答番号は採点の最小単位となる採点項目に対応する。以下、解答番号を用いて各項目を区別する。先に述べたとおり、第7問の問1は選択枝から二つの解答を選ぶ問題であり、項目13と14は解答の順序を問わない。また、どちらか一方のみが正答の場合には1点が与えられる。したがって、解答順序に関わらず、二つの正答のうち、特定の正答（選択枝番号が若い方の正答）を選択した者の割合を項目13の正答率

表4 各設問の正答率と識別力

問題番号	設問	解答番号	全体	
			正答率	識別力
第1問	1	1	0.39	0.43
	2	2	0.36	0.40
第2問	1	3	0.56	0.47
	2	4	0.09	0.24
第3問	1	5	0.54	0.48
	2	6	0.51	0.42
第4問	1	7	0.63	0.54
	2	8	0.40	0.45
	3	9	0.52	0.40
第5問		10	0.41	0.26
第6問	1	11	0.69	0.45
	2	12	0.35	0.42
第7問	1	13	0.64	0.48
	1	14	0.68	0.40
	2	15	0.59	0.32
第7問	3	16	0.25	0.35
		17	0.35	0.34

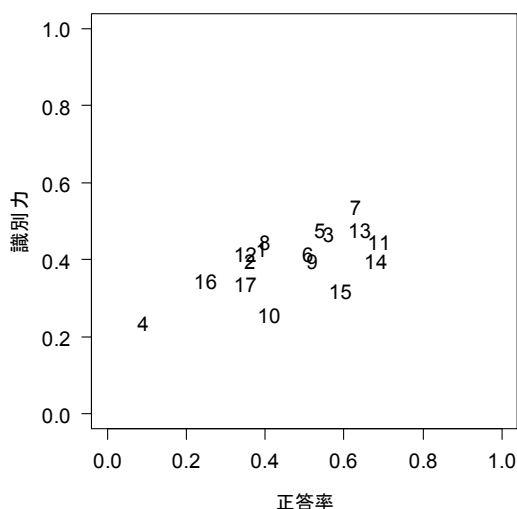


図4 正答率と識別力の分布

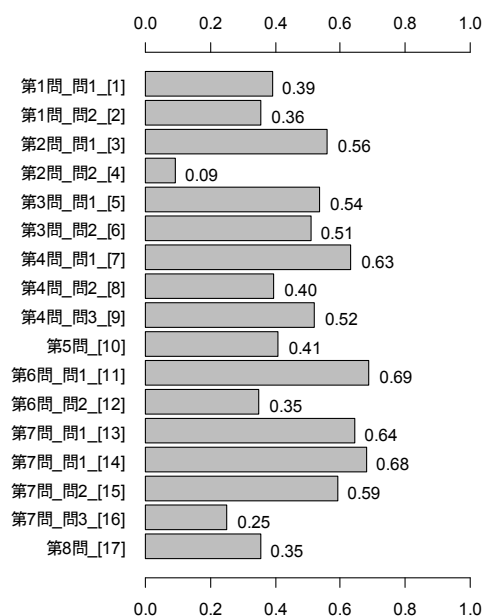


図5 各設問の正答率

欄に、もう一方の正答を選択した者の割合を項目14の正答率欄に示した。また、各設問の識別力はI-T相関係数（Item-Total Correlation Coefficient）によって表すこととした。

図4は、表4の正答率を横軸に、識別力を縦軸にとって各設問（採点項目）を解答番号で表して布置したもので、設問の正答率と識別力の分布を表す。項目4（第2問の問2）を除き、極端に正答率が低い設問や高い設問

は無く、0.3～0.7程度の正答率の設問が偏り無く出題されていることが分かる。また、正答率が0.6付近で識別力が高くなる傾向が見られる。

図5は設問順に正答率を棒グラフで表したものである。冊子の前半と後半で正答率の差があるといった傾向は特に見られないが、各大問内では、前にある設問より後にある設問の方が正答率が低い傾向が見られる。これは同じ問題設定の下にある設問で構成されている大問でも、全く独立な問題で構成されている大問でも同様であった。

4.3 主成分分析

各設問は実践的な言語運用力を測定するという共通の枠組みの中で作成されているが、表2に示したような素材文の言語や能力分類、あるいは読む文章の長さなどの違いがある。本節では、設問の独自性や共通性などの特徴を調べるために、各採点項目の正誤を表す17項目の解答データに対して相関行列を用いた主成分分析を行った結果を示す。

表5 各設問（採点項目）の主成分負荷量

問題番号	設問	解答番号	主成分			
			第1	第2	第3	第4
第1問	1	1	0.43	0.23	0.36	-0.20
	2	2	0.39	0.22	0.37	-0.28
第2問	1	3	0.49	-0.20	0.17	-0.26
	2	4	0.24	-0.23	0.22	0.51
第3問	1	5	0.50	-0.36	0.04	-0.06
	2	6	0.43	-0.45	0.11	-0.15
第4問	1	7	0.58	0.26	-0.11	-0.09
	2	8	0.46	0.36	-0.16	0.12
	3	9	0.38	0.43	0.13	0.02
第5問		10	0.16	0.39	0.16	0.29
第6問	1	11	0.47	-0.20	0.14	-0.07
	2	12	0.42	-0.25	0.03	0.32
第7問	1	13	0.51	-0.04	-0.23	-0.03
	1	14	0.40	0.07	-0.54	0.22
	2	15	0.28	0.09	-0.53	-0.36
	3	16	0.33	0.14	0.09	0.39
第8問		17	0.30	-0.24	-0.22	0.12
固有値			2.88	1.26	1.14	1.03
寄与率(%)			16.9	7.4	6.7	6.1

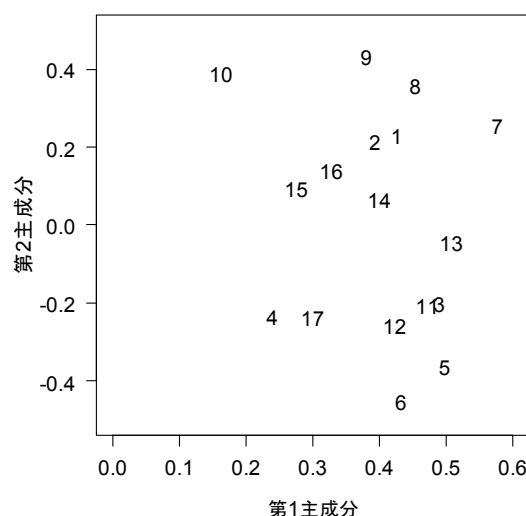


図6 各設問の主成分負荷量（第1-第2成分）

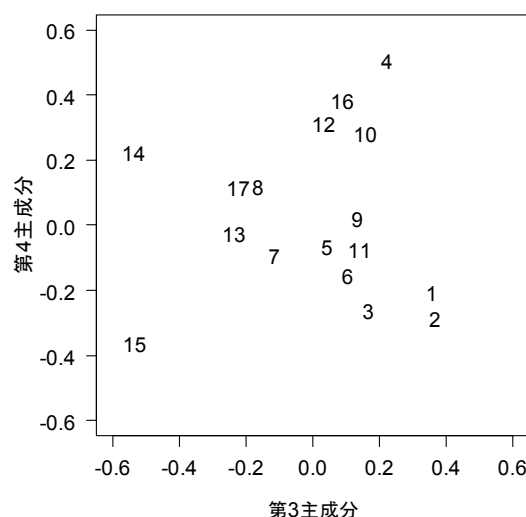


図7 各設問の主成分負荷量（第3-第4成分）

表5は各設問（採点項目）の第1～第4主成分の負荷量（主成分得点と個々の項目得点間の相関係数）、および主成分の固有値と寄与率を示したものである。寄与率は第4主成分までの合計で37.1%とそれほど大きくなく、第5主成分以降についても、4～6%の寄与率で緩やかに減少し、大きな変化はなかった。

図6は第1主成分を横軸に、第2主成分を縦軸にとった主成分負荷量プロットである。図中の数字は各設問（採点項目）の解答番号を表す。同様に、図7は第3主成分を横軸

に、第4主成分を縦軸にとった負荷量プロットである。

第1主成分は全て正であり、0.15～0.60の狭い範囲に分布している。このことから、第1主成分は言語運用力の総合的な能力を表す成分と考えられる。

第2主成分は、-0.1付近を境にして、正の方向に日本語問題（項目1, 2, 7, 8, 9, 10, 13, 14, 15, 16）、負の方向に英語問題（項目3, 4, 5, 6, 11, 12, 17）が明確に分かれて分布している。このことから第2主成分は素材文の言語の違い（日本語と英語の違い）を表す成分と考えられる。

第3主成分では、負の値をとっている項目は、項目17を除き、すべての項目（7, 8, 13, 14, 15）が表1の能力分類のL1（情報を正しく読み取る能力）だけに対応している。一方、正の値をとっている項目は、項目4を除き、すべての項目（1, 2, 3, 5, 6, 9, 10, 11, 12, 16）がL2（理解や解釈を行う能力）またはL3（推測、評価、判断等を行う能力）に対応している。このことから、第3主成分は情報の読み取りという比較的低下の能力と、読み取った情報に基づいて思考するという比較的高次の能力を分ける成分であることが推測される。これらの結果から表1の能力分類が一定の因子的妥当性を有していることが示唆される。

例外となった項目17はヨーロッパ旅行の体験を述べた100語程度の英文を読み、話し手がいくつの国を訪問したかを答える問題である。理解や解釈を行う能力（能力ラベルL2）に関する問題として作成されたものであるが、予想したほど文章間の関連がなく、読解できない部分が多少あったとしても、丁寧に訪問国を数えるだけで解答できたため、上記の主成分分析の結果において、情報を正しく読み取る能力（能力ラベルL1）に関す

る性質を示したものと思われる。

同じく例外となった項目4は二者による短い会話文（2往復の会話）を読み、美術館の入場料の合計額を答えさせる問題である。情報を正しく読み取る能力（能力ラベルL1）に関する問題として作成されたものであるが、予想したよりも文章の読解が難しく、会話の内容から入場料を決める条件を整理し、実際に入場料の合計額を計算するという作業が必要なため、上記の主成分分析の結果において、理解や解釈を行う能力（能力ラベルL2）や推測、評価、判断等を行う能力（能力ラベルL3）に関する性質を示したものと思われる。

さらに、図6、図7ともに同じ大問に属する項目どうしが比較的近くに位置する傾向が見られる。同じ大問内の項目どうしの関連性が、異なる大問の項目どうしの関連性よりも強いと考えられる。

4.4 大問間の相関

表6、7、8に大問間の相関係数を示す。表6は日本語問題で構成されている大問間、表7は英語問題で構成されている大問間、表8は日本語で構成されている大問と英語問題で構成されている大問間の相関係数である。

表6の相関係数の平均は0.18、表7の相関係数の平均は0.20、表8の相関係数の平均は0.14であり、同じ言語の大問間の相関係数よりも異なる言語間の相関係数の方が小さい値となった。また、日本語問題間よりも英語問題間の相関係数の方が大きい値となった。先に述べたとおり、英語問題は大学入試センター試験の「英語リスニング」の問題を利用しており、ある程度問題形式が類似していたことが影響している可能性があると思われる。

但し、項目数の少ない大問間の相関係数よりも項目数の多い大問間の相関係数の方が大きくなる傾向があると考えられることや、そもそも各大問に含まれる項目数が少ないこと

表6 日本語問題間の相関

	第4問	第5問	第7問
第1問	0.25	0.10	0.20
第4問		0.12	0.32
第5問			0.07
平均：0.18			

表7 英語問題間の相関

	第3問	第6問	第8問
第2問	0.28	0.25	0.11
第3問		0.29	0.14
第6問			0.14
平均：0.20			

表8 日本語問題－英語問題の相関

	第2問	第3問	第6問	第8問
第1問	0.19	0.16	0.19	0.08
第4問	0.21	0.20	0.22	0.10
第5問	0.04	0.00	0.07	0.01
第7問	0.20	0.23	0.24	0.15
平均：0.14				

などから、上記の結果はあくまで参考程度の理解にとどめることが適当と考えられる。今後、より詳細に調べるためには、項目数に注意することに加え、大学ごとの相関係数の違いなど、下位集団に着目した分析も重要と考えられる。

5 まとめと考察

今回、2000名以上の学生を対象に「言語運用力」試験のモニター調査を行い、試験全体および個々の設問の特性を調べることができた。その結果、概ね意図したとおりの困難度と識別力を有する試験問題が作成できたと思われる。

今回の調査で実施した「言語運用力」試験は、AO入試や推薦入試の受験者など、従来の共通第一次学力試験および大学入試センター試験では主な対象となっていなかった幅広い学力層の基礎的学力を診断・評価することを目的としたものであり、当面は現状のAO入試、推薦入試、大学入試センター試験と同

時期に実施する選抜試験を想定したものである。

しかしながら、近年、競争率の高い一部の大学を除けば、入学試験の選抜試験的性格が薄れ、事実上、資格試験や入学後の学修の参考とする評価試験のような性格も帯びつつあることに鑑みれば、今回の「言語運用力」試験の資格試験的、評価試験的な応用について検討することも有意義であると思われる。

今回、全体の正答率は5割弱であったが、この試験を大学入学にあたっての最低限必要な言語運用力を確認する試験として用いるのであれば、全体の正答率を0.6～0.7程度、またはそれ以上に設定することも考えられる。また、今回はどのような試験問題がどのような受験者応答を生じさせるかを調べるためにある程度正答率にバラツキ（0.3～0.7程度）を持たせたが、正答率の高い問題から低い問題までを一様に出題せず、0.6～0.7程度の正答率の設問を中心に問題冊子を構成することも考えられる。図4では、ちょうどこの辺りの正答率で高い識別力を示しており、識別力の観点からも妥当な問題構成になるとと思われる。

また、項目の正誤データを主成分分析した結果では、項目17と項目4が問題作成時に意図した能力分類とは異なる能力に対応していることが示唆されたが、4.3節において問題内容を精査した結果を考慮すると、むしろ主成分分析の結果から示唆される能力分類の方が妥当であったと思われる。

4.3節でも述べたが、主成分負荷量のプロットにおいて、L1に代表される“情報の読取り”という比較的低次の能力とL2とL3に代表される“読み取った情報に基づいて思考する”という比較的高次の能力を分ける成分が確認された。

一方、L2とL3を明確に分ける成分は見られなかった。原因としては、L2とL3の能力がそもそも概念的に分けられない可能性

があることや、L2とL3の能力に精確に対応した操作を個々の問題に組み込むことができなかつた可能性などが考えられる。さらに、第2、第3主成分の寄与に対して、L2とL3を分ける成分の寄与が小さすぎた可能性も考えられる。この場合、言語を日本語に統一したり、L1に対応する問題を除外したりすることによって、L2とL3の違いがより明確になる可能性があると思われる。

今後は「言語運用力」試験が意図したとおりの能力をどの程度測定しているか、また実際の受験者の応答において、試験の測定している構成概念がどのような構造として現れるかについてさらに分析を進めて行く予定である。

例えば、教科科目固有の知識に依らない総合試験を冊子レベルで分析した研究（伊藤ほか、2010）では、試験の妥当性を調べるために、総合試験と教科科目の学科試験の得点に対して因子分析を行い、得られた因子と受験者の能力・資質に関するアンケートとの関係性を調べる手法が用いられている。また、項目レベルの分析としては、法科大学院適性試験の各設問の正誤データに対して因子分析を行い、適性試験で測定される能力の構造を検討した研究も行われている（椎名ほか、2011）。

今回の調査でも基本的な能力・資質に関するアンケートを実施しており、上記のような因子分析的手法を用いて、試験成績とアンケートの関連について分析を進めていくことが有効と思われる。

参考文献

伊藤圭・林篤裕・椎名久美子・田栗正章・小牧研一郎・柳井晴夫(2010).「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・宮埜寿夫・椎名久美子・荒井清佳・

桜井裕仁・田栗正章・小牧研一郎・安野史子(2014).「大学入学志願者の基礎的学力測定のための英語問題の試作とモニター調査による予備的検討—正答率分析図を用いた問題内容と受験者の応答の事例分析—」『大学入試研究ジャーナル』24, 59-67.

文部科学省 平成25年度国公立大学・短期大学入学者選抜実施状況の概要 2013年10月18日

< http://www.mext.go.jp/b_menu/houdou/25/10/1340441.htm >
(2014年4月25日).

桜井裕仁・田栗正章・安野史子・小牧研一郎・荒井清佳・伊藤圭・椎名久美子・宮埜寿夫(2014).「大学入学志願者の基礎的学力測定のための「数理分析力」の調査とその予備的検討」『大学入試研究ジャーナル』24, 51-58.

桜井裕仁・田栗正章・安野史子・小牧研一郎・荒井清佳・伊藤圭・椎名久美子・宮埜寿夫(2015).「「数理分析力」試験の開発と検討—大学新入生に対する2013年度調査の結果とその分析—」『大学入試研究ジャーナル』25, 21-28.

椎名久美子・荒井清佳・小牧研一郎(2011).「法科大学院適性試験の問題項目の特性—因子分析による検討—」『日本テスト学会第9回大会発表論文抄録集』92-93.

椎名久美子・宮埜寿夫・伊藤圭・荒井清佳・桜井裕仁・小牧研一郎・田栗正章・安野史子(2014).「大学入学志願者の基礎的学力測定のための枠組みの検討および「言語運用力」についての予備的分析」『大学入試研究ジャーナル』24, 41-49.