

名古屋大学教育学部における論述的学力検査

小嶋秀夫・村上 隆（名古屋大学教育学部）

1. 論述的学力検査の特質とその導入

論述的学力検査は、名古屋大学教育学部において3年間の論議と準備を経て、1985年度から導入された第2次試験の科目（試験時間は2時間）である。募集要項においては、「高等学校の国語・社会の学習を前提とする」、「与えられた素材を読み解き、また、分析、総合し表現することを含み、国語Ⅰ、国語Ⅱ、国語表現、現代の学習を基礎とした応用力をためすものである」とされている。後に示す内容の概略からも分かるように、与えられる素材は社会科学・人文科学の広い範囲から引き出される。

この論述的学力検査は、入試科目や入試研究の大分類では「小論文」に含まれているが、導入当時の典型的な小論文とは明らかに違った形式と内容をもつ試験を目指したことは、その独自の命名にも現れている。

すなわち、それはまず、ある側面の学力を測ろうとするものであって、態度・動機・経験などを調べるものではない。そしてそれは、当時の共通第1次学力試験のような多肢選択法によらない記述式の試験である。またそれは、短い課題を提示して、かなり長いエッセイを求める小論文（課題型とも呼ばれる作文タイプのもの）とは別のものを狙っている。

そのような課題型の小論文に対して、近年全国的に増加してきたのが資料型の小論文であり、論述的学力検査はそのカテゴリーに含められよう。ただし、論述的学力検査は1つの大問からなるのではなくて、異なる素材を用いた大問を複数（3～4）個出題し、それぞれの大問はいくつかの小問からなるという構成になっている。1つの問いに1つの答えを求めるという集束的産出（convergent production）形式の他に、質のよい多様な答えを出すという発散的産出（divergent production）を求めることもある。

このような選択は、理論的な検討と、先行していた他大学・他学部の試みの評価とに基づいて行われた。その後のわれわれ自身の経験から、出題内容にも若干の変化が認められ

るが、この科目の基本的な狙いと様式とは、現在まで一貫している。

2. 論述的学力検査の実施・研究体制

実施体制

名古屋大学教育学部では、論述的学力検査に対して、導入の最初の時期から、検査の実施とその結果を分析する研究とを有機的に関連づけた体制で臨むことにしていた。その体制は、過去6年の内に徐々に整備され、現在7年目を迎えようとしている。学力検査委員・出題委員・採点委員のチームと、学部の入学試験方法検討委員会とは、別々の組織であるが、一部のメンバーが両方に属することによって、両者間の連携を図るとともに、情報のフィードバックと結果の評価を容易にした。

採点に関しては、出題委員のチーム（問題の校閲を依頼する委員を除く）により行うことが最初から決められていた。出題の意図を徹底し採点の基準を共有するためには、それが最善の方法であると考えられたからである。それは検査の信頼性を確保するために必要な条件だと考えられた。もちろん、出題委員を兼ねない採点委員を訓練することも可能であったが、そのための訓練時間を費やすよりも、出題委員だけで採点してしまう方が手っ取り早いと判断されたのである。また、方式の継続性を確保するために、委員の任期を長くし、部分的に交代する方式をとっている。また、ある小問については、全受験者（学生の募集単位は1つである）の答案を通して、同一の組の採点者が採点するという原則は、当初から今日まで守っている。それは受験者に対する公平性を確保するためのものであるが、同時に、複数の採点者による採点結果の相関を低める誤差要因と、論述的学力検査と他の変数（入学試験および入学後の成績など）との連関を低める誤差要因とを少なくする効果をもつと理論的に考えられる。

検査の信頼性（reliability）の検討

さて、論述的学力検査導入後の初期の段階では、検査の信頼性に関する情報として、採点者間の一致度（相関）と問題間あるいは小問間の相関に関心が向けられた。入学試験に利用される他の情報（共通第1次学力試験、第2次学力試験の他の教科、および高校調査書）と論述的学力検査との関係を調べるにしても、まず論述的学力検査の信頼性がどの程

度であるかを確かめることが先決であったからである。

最初の2年(1985, 86年度)は、受験機会の複数化以前の時期で、受験者数が多くなかったため、すべての受験者のすべての小問に対する解答を、4人が別々に採点した。この時期には、採点に先だって基準を統一する努力を払った後は、採点結果を調整することはせず、個々の受験者の各小問に対する4人の採点結果の平均値を、そのまま得点として使用した。そして、採点の信頼性に関する情報は、事後の分析にのみ使用された。その結果、後に報告されるように、採点者間の一致度は概して高いことが分かるとともに、特異的に起こる一致度の低下は、解答の仕方を含んだ出題の工夫でかなりの程度まで防げる見通しが立った。われわれは採点者間の一致度として、専ら相関係数を指標としている。複数の採点者の間で、平均値と標準偏差とを一致させることは困難である。その違いは、受験生に対する不公平を招かない。ただし、分散の大きい採点者が、論述的学力検査の総得点に相対的に大きな寄与をすることになる。

上記の結果を踏まえて、受験機会の複数化によって受験者数が激増した3年目(87年度)からは、1つの小問に対する採点者の数を2名にした。後にも述べるように、そのことによって、採点者間の相関や問題間の相関のパターンに目立った変化は起こらなかった。しかし、得点の決定に4人のデータを用いるのと2人のそれによるのとでは、1人の採点者が冒すエラーの影響度が違う。

そこで4年目(1988年度)からは、2名1組の採点者が採点基準を合議してから独立的に行った小問毎の採点結果(平均値、標準偏差、2人の採点の相関係数と、大きな食い違いがあったケースの表示)を両者にフィードバックすることにした。その情報に基づいて、採点者は自分の採点結果を見直しをする。これは、2人が歩み寄って調整するのではなく、それぞれが自分の採点が適切であるかを見直す手続きである。そして個々の採点者が、必要と認めた場合には得点を修正する。

「大きな食い違い」を得点の差の絶対値で決めることは、平均値と標準偏差が採点者間で同じではないので、合理的ではない。そこで、1989年度からは、採点者毎に標準得点を求め、その差が一定の範囲を越えた場合に、マークをつけて採点者に警告を発し、散布図(scatter diagram)を示すことにしている。警告がついたケースで、一方または両方の

採点者が得点を修正する場合には、最初につけた得点は誤りだということになる。そこで、誤りの原因が何であったかを分類カテゴリーによりコード化して、自己申告することにした。後に触れるように、複数の基準を適用せざるをえない質的な採点では、どうしても基準の適用の仕方を誤ることが起こる。上記のような見直しによって、採点者間の相関はほぼ満足すべき水準に達する。

なお、細かい技術上の問題として、丸めの誤差のことに触れておく。2名の採点結果を平均して最終得点とする過程で、4捨5入をすると、受験者は0.5点の得をすることになる。2名がつけた得点の合計が奇数となるか偶数となるかはランダムに起こると考えられるが、実際にやって見ると、総得点に関して、恩恵に浴する機会のない受験生から数点分の得をするものまでに分かれる。これは無視できない差であるので、ある小問で得をした場合には、後続する小問で次にその機会があったときには切捨てをすることによって相殺する(+0.5-0.5)処理を続けることにしている(最大でも0.5点しか得をしない)。

以上のようなフィードバックや処理を正確かつ迅速にするためには、計算機を必要とする。最初の小問毎の得点の入力ミスを防ぐために、2人で別々に入力した結果を計算機で照合するところから、これまで説明した各種の分析まで、採点作業と計算とが交互に行われる。少なくとも論述的学力検査の信頼性と内部構造の統計的性質については、最終採点結果が出るのと同時に判明する。換言すると、採点作業と入試研究とが同時に進行するのである。通常の採点の一致度の研究は、入試の過程の外で模擬的に行われる。名古屋大学教育学部の論述的学力検査に関しては、比喩的にいうと、*in vitro*ではなくて*in vivo*の入試研究をして入試方法の改善に努力しているところに特徴がある。それを可能にしたのは、新しく導入した試験科目を、公平で正確なものにしようという関係者の熱意と協力的体制であった。また受験機会の複数化以来、名古屋大学教育学部は一貫して連続方式のA期間に留まり続けたことも大きい要因である。最大でも300名以下の受験生であるが、きめ細かい出題採点の手続きを取るためには、現在の程度にゆとりのある期間が必要である。

検査の妥当性 (validity) の検討

次に、論述的学力検査は、それがとらえようとした受験生の特徴を実際にどれほど測定しているのが重要な研究課題となる。そこでまず、論述的学力検査が、入学試験におけ

る他の科目（共通第1次学力試験，第2次試験の他の科目）および高等学校の調査書とどのように関連しているかを調べた。妥当性を推測する資料としては，きわめて間接的なものであるが，当初にえられるのは，この情報に限られていた。これに関して，現在までに6年分の結果が出ている。

次に，入学者の学年進行に伴って，入学後の成績が蓄積される。入学者の内部において，高い順位で入学したか・低い順位で入ったかということと，入学後の成績との間に，なんらかの関連があるかどうかの検討は，検査の予測的妥当性（predictive validity）に関する重要な情報を提供する。

幸いにも名古屋大学では，1985年度以降の入学者について，教養部2年終了時点での成績を確保して入学試験の追跡研究をすることが，全学部について可能となった（名古屋大学・入学者選抜方法研究委員会報告書〔昭和61，62，63年度合併〕，1990年）。これに関しては，これまでに4年分の結果が出ている。最後に，学部進学後の成績がこれまで2年分集まっている。これらの結果は，後に説明される。

3. 過去の論述的学力検査の内容

以下に示すのが，これまでの試験問題の概要である。

年度	番 号	内 容	出 典
8 5	[1]	オランダの海外進出が鎖国下にあった日本に与えた影響を題材とした文：語句説明（50字2問），空所補充，要約（100字）	「世界の歴史8」 大野真弓
	[2]	エジプトのピラミッド建設とアポロ計画について述べた文：書き取り（5問），ピラミッド建設とアポロ計画の類似点（400字）	「技術と人間」 星野芳郎
	*[3]	明治政府が行った国民教化政策に応えた新聞報道と僧侶・神官にスポットをあて，それらの内容上	「明治メディア考」 加藤秀俊・前田 愛

	<p>のずれを論じた対談：明治初期のコミュニケーションの状況の説明〈3つの視点を指示〉(400字)</p> <p>*[4] 1)19世紀イギリスの教育普及による国民の読書欲の向上, 2)ロイター通信の起源・沿革の2つの文と年表：19世紀のイギリスの公衆の関心の変化を述べる(400字) <年表></p>	
8 6	<p>[1] 「進歩の観念」という観点からヨーロッパ文化史(ギリシア・ローマ文化, ヨーロッパ中世の文化, ルネサンス, 古典主義, 啓蒙主義など)について述べた文：書き取り(5問), 個所説明(100字2問), 内容説明(200字)</p> <p>[2] ノーベル科学賞受賞者数の主要国別の比較資料：資料から読み取れること(200字), 諸外国と比較して, わが国の受賞者が少ないことが, 日本人の創造性の低さの現れであるという命題の妥当性を論述する(200字) <表></p> <p>[3] カリフォルニアにおける自動車の普及と映画産業の発展を述べた文：理由説明(200字), 内容説明(200字)</p> <p>[4] ルビの廃止という国語政策の利点・欠点を述べた2つの文：2つの資料文の見解を「教育」の観点を中心にそれぞれ要約(100字2問), 対立した2つの立場から, 同じ図が違って解釈されることを説明する(200字2問) <図表></p>	<p>「ヨーロッパからの発想」 木村尚三郎</p> <p>「続イギリスと日本—その国民性と社会」 森嶋通夫</p> <p>「カリフォルニア・ストーリー」 石川好</p> <p>「山本有三先生と国語運動」松平忠則, 「国語改革の歴史(戦前)」大野晋, 「図説日本語」林大監修</p>
	<p>[1] ルネサンス時代の人間の「高貴さ」の条件の一つ</p>	<p>「ルネッサンス夜話」</p>

		としての「財産」を論じた文：書き取り（5問） ， 個所説明(100字)， 要旨(200字)	高階秀爾
	[2]	新技術開発の3つの段階と， 経済システムの変遷 に関して述べた文：図の内容説明(300字)， 個所 説明(100字) <図>	「都市と情報」 佐貫利雄
8 7	[3]	現今の風潮である歴史に対しての「実感信仰」と 「擬現代化」を批判する文：箇所から考えられる 事実とその概略・教訓を述べる(200字)， 内容説 明(200字)， 筆者の見解のまとめ(200字)	「近代合理主義の光と 影」 竹内啓
	[4]	日本の鎖国政策に関するアメリカと日本における 批判文：当時の両国の批判論の共通点と相違点 (200字)， その批判論がその後の両国の行動と日 米関係に及ぼした影響の記述(200字)	「日米の鎖国批判論」 M. W. スティール
	[1]	アメリカの都市におけるストリートの役割と2つ のタイプのストリートについて述べた文：書き取 り（5問）， 個所説明(200字)， 内容説明（300 字）	「カルチャー -マス- カルチャー」 青木保
	[2]	明治の思想傾向について論じた文：内容説明(200 字2問)， 個所説明(200字)	「戦中と戦後の間」 丸山真男
8 8	[3]	歴史人口学の手法を用いて作成されたドイツの特 定の地域における特定の期間中の死亡者の年齢別 構成比のグラフと， 17～20世紀後半のドイツの平 均余命の変化のグラフ：グラフからの読み取り (100字)， 変化をもたらした要因を3つまで説明， 「死」に対する個人の意識の変化の推測(200字) <グラフ>	

89	[1]	D. デフォウとJ. スウィフトによる当時のイギリス社会の見方の違いについて述べた文：両者の違いの説明(300字)， 個所説明(200字)	「社会科学の方法」 大塚久雄
	[2]	日本人の「精神的鎖国」状況の打開について論じた文：書き取り(5問)， 個所説明(100字)， 内容説明(200字)	「存亡の条件」 山本七平
	[3]	産業化・教育費・就学率および出生率に関する設問：産業の発展と教育費増大との関係の理由説明(200字)， 経済的発展と出生率低下の関係を考えて子どもを産む効用と不効用を各3点に分けて論じる， 就学率と出生率の変化を示した図から， 前者の向上と後者の低下との関連を解釈(グラフ)	
90	[1]	日本の近代化におけるナショナリズムのあり方について論じた文：書き取り(5問)， 抽象的に述べられた内容例を3つまで列挙して説明， 文中の中心的概念の説明(200字)	「西欧の衝撃と日本」 平川祐弘
	[2]	世界の森林面積と日本の木材輸入量の推移を表すグラフと表：表からの読み取り(200字)， 森林面積の減少に関する理由説明3つ(グラフ・表)	
	[3]	「言」と「行」の意味・関係を論じ， 両者の遊離による現代マスコミ界の言論のあり方を批判した文：2つの概念の意味と関係の説明(200字)， 所説明(200字)， 筆者の主張(200字)	「モラトリアム人間の時代」 小此木啓吾

(注) 85年度：〔1〕〔2〕は必答， *印〔3〕〔4〕はいずれか一方を選択。

4. データ分析の結果と考察

分析の対象および目的

前述のように、この報告で分析の対象としたのは、1985～1990年度の6年度分（教養課程の成績は85～88年度入学者、専門課程の成績等は85、86年度入学者のみ）である。この間の受験者（2次試験全科目）と合格者の数は、以下の通りである。

	受験者	合格者	入学者
1985年度	113	65（追加合格含む）	60
1986年度	110	73	71
1987年度	285	87（追加合格含む）	70
1988年度	282	92	72
1989年度	280	99	72
1990年度	214	81	77

これも前述したように、以下の分析は、論述的学力検査の信頼性と妥当性について検討することを目的となされた。分析の視点を改めて述べれば、

- (1) 採点者間の一致度、および、問題間の相関にもとづく、信頼性の検討。すなわち、論述的学力検査の内部構造の分析である。
- (2) 入学試験の他の科目、および調査書の成績と、論述的学力検査との関係、選抜資料の中における、論述的学力検査の位置の検討である。
- (3) 入学後の成績を基準とした、妥当性の検討。(2)とあわせて、論述的学力検査が測定している能力の内容を明らかにすることを目的とする。

以下、これらの視点ごとに、得られた結果を略述する。

論述的学力検査の信頼性 1——採点者間の一致度

論述的学力検査においては、2つの側面から信頼性を検討することができる。第1は、採点者間の一致度である。表1～7に86～90年度の各小問ごとの採点者間相関（相関行列の対角要素の（ ）内の数値）を示した。85、86年度においては、各小問を4人の採点者が採点したために、すべての採点者の組み合わせである6個の相関係数が得られるが、表に示したのはその平均値である。また、88年度以降、採点結果を採点者にフィードバックし、不一致の大きい場合について再度協議する等の方法を取り入れたため、89、90年度については、調整の前後の結果をともに示している。なお、88年度については、調整後の結果のみである。これらの表には、各小問及び総点の平均、標準偏差（SD）、

表1 86年度の小問間相関と採点者間一致度

漢字	I-2	I-3	I-4	II-1	II-2	III-1	III-2	IV-1	IV-2	IV-3	IV-4	総点	
漢字	1.00												
I-2	-0.02	(0.51)											
I-3	0.19	0.17	(0.45)										
I-4	0.25	0.29	0.21	(0.62)									
II-1	0.10	0.21	0.17	0.30	(0.72)								
II-2	0.10	0.04	0.06	0.18	0.63	(0.69)							
III-1	-0.01	0.30	0.31	0.16	0.20	0.08	(0.45)						
III-2	0.03	0.13	-0.04	0.19	0.02	-0.05	0.11	(0.83)					
IV-1	0.17	0.13	0.16	0.56	0.25	0.20	0.18	0.11	(0.49)				
IV-2	0.06	0.12	0.24	0.41	0.33	0.23	0.16	0.00	0.64	(0.61)			
IV-3	0.13	0.06	0.15	0.19	0.17	0.13	-0.01	0.25	0.15	0.14	(0.74)		
IV-4	0.13	0.15	0.16	0.27	0.19	0.16	0.12	0.20	0.38	0.50	0.62	(0.63)	
総点	0.24	0.36	0.31	0.59	0.61	0.50	0.39	0.53	0.52	0.51	0.57	0.63	1.00
満点	10	10	10	20	25	25	25	25	10	10	15	15	200
平均	9.3	6.1	7.4	11.6	12.7	12.3	20.1	14.3	7.0	7.4	7.2	8.7	124.2
SD	1.02	1.52	1.21	2.61	3.75	3.66	2.61	6.36	1.21	1.5	3.52	2.89	16.45
最高	10	8	9	16	19	19	24	23	9	9	12	13	158
最低	6	0	1	0	0	0	8	0	0	0	0	0	45

表2 87年度小問間相関と採点者間一致度

漢字	I-2	I-3	III-1	III-2	III-3	II-1	II-2	IV-1	IV-2	総点	
漢字	1.00										
I-2	0.22	(0.73)									
I-3	0.16	0.27	(0.55)								
III-1	0.08	0.08	0.13	(0.53)							
III-2	0.11	0.12	0.18	0.11	(0.63)						
III-3	0.16	0.22	0.25	0.16	0.33	(0.39)					
II-1	0.08	0.09	0.07	0.07	0.00	0.01	(0.71)				
II-2	0.08	0.04	0.07	-0.04	0.20	0.14	0.30	(0.57)			
IV-1	0.04	0.05	0.05	0.13	0.07	0.09	0.02	0.01	(0.60)		
IV-2	0.11	0.16	0.11	0.21	0.15	0.14	0.17	0.08	0.38	(0.61)	
総点	0.34	0.43	0.48	0.35	0.55	0.53	0.48	0.39	0.39	0.59	1.00
満点	10	15	25	10	20	20	35	15	25	25	200
平均	7.2	10.2	16.1	3.1	10.3	7.9	22.6	8.4	15.5	9.1	110.2
SD	2.25	3.10	4.24	2.48	6.02	4.12	6.52	2.91	3.83	5.41	19.4
最高	10	15	25	9.5	19	18.5	33	13	22.5	20	156
最低	0	0	0	0	0	0	0	0	0	0	38

表3 88年度小問間相関と採点者間一致度

漢字	I-2	I-3	II-1	II-2	II-3	III-1	III-2	III-3	III-4	総点	
漢字	1.00										
I-2	0.10	(0.05)									
I-3	0.11	0.19	(0.61)								
III-1	0.07	0.01	0.13	(0.83)							
III-2	0.05	0.11	0.11	0.13	(0.75)						
III-3	0.04	0.17	0.07	0.18	0.33	(0.84)					
II-1	0.07	0.01	0.03	0.02	-0.00	0.03	(0.83)				
II-2	0.14	0.03	0.01	0.09	0.10	0.08	0.50	(0.77)			
IV-1	0.11	-0.05	0.00	0.14	0.07	0.09	0.19	0.34	(0.87)		
IV-2	0.07	-0.07	0.11	0.11	0.22	0.09	0.20	0.42	0.37	(0.89)	
総点	0.27	0.17	0.28	0.46	0.50	0.49	0.41	0.61	0.54	0.66	1.00
満点	10	20	20	20	20	20	20	27	18	25	200
平均	7.0	12.3	11.0	5.8	7.4	6.4	10.8	12.9	11.5	12.6	97.7
SD	2.08	1.55	2.43	5.08	4.60	4.98	3.20	4.40	4.14	6.10	18.85
最高	10	16	16.5	20	17	15	20	26.5	18	22.5	148
最低	0	7.5	0	0	0	0	0	0	0	0	34.5

表4 89年度小問間相関と採点者間一致度(調整前)

漢字	I-1	I-2	II-2	II-3	III-1	III-2-1	III-2-2	III-3	総点	
漢字	1.00									
I-1	0.13	(0.70)								
I-2	0.04	0.30	(0.69)							
II-2	0.10	0.00	0.06	(0.32)						
II-3	0.07	0.31	0.23	0.10	(0.76)					
III-1	0.09	0.16	0.00	0.06	0.14	(0.68)				
III-2-1	-0.04	0.20	0.15	0.09	0.19	0.11	(0.84)			
III-2-2	0.13	0.14	0.16	0.09	0.23	0.17	0.27	(0.77)		
III-3	-0.02	0.10	0.17	0.08	0.19	0.15	0.21	0.27	(0.80)	
総点	0.23	0.52	0.49	0.25	0.63	0.45	0.45	0.53	0.65	1.00
満点	10	30	30	20	30	20	15	15	30	200
平均	7.6	22.5	16.2	14.8	18.6	11.0	10.5	10.5	12.5	124.1
SD	2.16	3.45	3.78	1.93	5.26	4.06	2.52	2.91	6.83	16.80
最高	10	27.5	24.5	19	28	19.5	15	15	27	156.5
最低	0	0	0	8	0	0	0	0	0	34

表5 89年度小問間相関と採点者間一致度(調整後)

漢字	I-1	I-2	II-2	II-3	III-1	III-2-1	III-2-2	III-3	総点	
漢字	1.00									
I-1	0.13	(0.83)								
I-2	0.05	0.28	(0.82)							
II-2	0.13	0.02	0.06	(0.75)						
II-3	0.07	0.29	0.23	0.13	(0.86)					
III-1	0.10	0.15	-0.01	0.08	0.14	(0.86)				
III-2-1	-0.03	0.18	0.14	0.10	0.21	0.12	(0.91)			
III-2-2	0.13	0.13	0.15	0.10	0.25	0.17	0.29	(0.90)		
III-3	-0.03	0.09	0.17	0.05	0.20	0.15	0.22	0.26	(0.90)	
総点	0.23	0.50	0.48	0.27	0.64	0.44	0.46	0.53	0.64	1.00
満点	10	30	30	20	30	20	15	15	30	200
平均	7.6	22.4	16.2	14.7	18.5	11.0	10.4	10.5	12.4	123.7
SD	2.16	3.54	3.84	2.01	5.31	4.15	2.54	3.01	6.88	17.05
最高	10	27.5	25	19	28	19.5	15	15	27.5	156.5
最低	0	0	0	7	0	0	0	0	0	34

表6 90年度小問間相関と採点者間一致度(調整前)

漢字	I-2	I-3	II-1	II-2	III-1	III-2	III-3	総点	
漢字	1.00								
I-2	0.09	(0.58)							
I-3	0.09	0.24	(0.47)						
II-1	0.08	-0.03	-0.05	(0.65)					
II-2	-0.00	0.14	0.06	0.08	(0.79)				
III-1	0.13	0.17	0.17	-0.02	-0.07	(0.59)			
III-2	0.19	0.16	0.11	-0.10	0.07	0.22	(0.56)		
III-3	-0.02	-0.03	0.01	0.08	0.03	0.12	0.24	(0.73)	
総点	0.29	0.52	0.40	0.21	0.53	0.47	0.52	0.40	
満点	10	20	20	25	50	25	25	25	200
平均	6.9	5.5	8.9	11.6	19.9	14.8	13.1	12.8	93.5
SD	2.02	4.44	3.23	2.96	7.29	4.37	5.14	3.93	14.94
最高	10	20	18	21	42.5	23.5	22.5	18.5	140.5
最低	2	0	0	0	0	0	0	0	42.5

表7 90年度の小問間相関と採点者間一致度(調整後)

漢字	I-2	I-3	II-1	II-2	III-1	III-2	III-3	総点	
漢字	1.00								
I-2	0.10	(0.84)							
I-3	0.10	0.21	(0.76)						
II-1	0.09	0.01	-0.03	(0.82)					
II-2	-0.01	0.18	0.03	0.07	(0.85)				
III-1	0.12	0.19	0.16	-0.03	-0.06	(0.79)			
III-2	0.17	0.14	0.08	-0.09	0.05	0.16	(0.71)		
III-3	-0.02	-0.01	0.01	0.09	0.04	0.12	0.25	(0.85)	
総点	0.29	0.56	0.38	0.22	0.53	0.46	0.49	0.42	
満点	10	20	20	25	50	25	25	25	200
平均	6.9	5.6	8.7	11.6	19.8	14.5	12.7	12.7	92.5
SD	2.02	4.76	3.31	3.03	7.29	4.50	5.17	4.03	15.29
最高	10	20	18	21	42.5	22	22.5	18.5	140.5
最低	2	0	0	0	0	0	0	0	41.5

最高点、最低点も示している。なお、「漢字」と記された変数は、毎年出題されている漢字の書き取りの問題であって、この部分は論述ではない。

また、表 8～15 は、85 年度も含めて、大問間の相関と採点者間相関を示したものである。85 年度は、ほとんどの大問が小問に分かれていないので、こちらの方に示した。85 年度の「知識」という変数は、この年のみ出題された歴史的事項に関する「穴埋め」問題である。

一部の例外を除き、採点者間一致度という意味での信頼性は高いと見て良いであろう。詳細に見ると、各小問の字数を減らした上で基準の設定をより緊密に行うようにした 86 年度に大きな上昇があり、採点者に結果のフィードバックを開始した 88 年度にもう一度大きな上昇がある。88 年度の I-2 は例外的に相関が極端に低い。これは SD が小さいことからわかるように、実際に答案の間にほとんど差がなかったことが理由であろう。この問いは論述形式の問いとしては最初に置かれたものであり、全員が着手するからますます差はつきにくくなるのであろう。

採点者間の不一致の原因の大部分は、採点者間の基準の不一致によるよりは、答案の誤読と基準の適用の誤りである。

この種の「主観」テストに対して、採点者による結果の不一致がしばしば心配されるが、論述的学力検査に関する限り、この点にはほとんど問題がないと結論してよからう。

論述的学力検査の信頼性 2 —— 問題間の相関

信頼性の第 2 の側面は、小問間、問題間の相関の高さである。もし、「論述」が、全体として単一の能力を測定しているとすれば、この相関が低いことは、個々の小問または問題の得点の誤差が大きいことを意味するとも考えられる。たとえば、受験者に 30 枚ずつコインを持たせ、それらを 10 枚ずつ同時に投げさせて、それぞれの回で表が出た枚数をその受験者の得点とする、といったテストを考えよう。これは、各 10 点満点の問題 3 つからなるテストであるとみなすことができる。枚数の判定には、2 人の判定員が付くものとする。このとき、2 名の判定員の結果はほとんどの場合一致するであろうから、たとえば表 15 の形の相関行列の () 内の値 (対角要素) は、ほぼ 1.00 となるであろう。しかしながら、各「問題」の得点は完全な偶然によって決まるから、問題間の相関はほぼ 0.00

表 8 85年度問題間相関と採点者間一致度

	漢字	知識	I-1	I-3	II-2	III*	IV*	総点
漢字	1.00							
知識	0.03	1.00						
I-1	-0.02	-0.03	(0.35)					
I-3	0.07	0.11	-0.10	(0.36)				
II-2	0.19	0.24	0.10	0.11	(0.28)			
III*	0.39	0.18	0.18	-0.02	-0.12	(0.26)		
IV*	0.00	0.02	0.18	0.08	0.28	*	(0.28)	
総点	0.46	0.45	0.25	0.46	0.60	0.63	0.68	1.00
満点	20	10	20	40	50	60	60	200
平均	18.8	4.3	13.2	21.4	31.0	36.9	39.3	126.6
SD	2.40	2.29	1.74	4.03	4.23	5.50	4.87	10.54
最高	20	10	18	28	39	46	49	147
最低	8	0	9	0	20	20	28	82

*IIIとIVは、どちらか一方を選択

表 9 86年度問題間相関と採点者間一致度

	漢字	I	II	III	IV	総点
漢字	1.00					
I	0.22	(0.64)				
II	0.11	0.27	(0.79)			
III	0.03	0.27	0.04	(0.74)		
IV	0.16	0.39	0.28	0.24	(0.74)	
総点	0.24	0.64	0.61	0.62	0.74	1.00
満点	10	40	50	50	50	200
平均	9.3	25.1	25.0	34.5	30.3	124.2
SD	1.02	3.86	6.69	7.14	7.01	16.45
最高	10	32	35	47	41	158
最低	6	6	0	14	4	45

表 10 87年度問題間相関と採点者間一致度

	漢字	I	III	II	IV	総点
漢字	1.00					
I	0.23	(0.63)				
III	0.16	0.30	(0.59)			
II	0.10	0.11	0.09	(0.70)		
IV	0.10	0.14	0.22	0.13	(0.65)	
総点	0.34	0.57	0.70	0.54	0.61	1.00
満点	10	40	50	50	50	200
平均	7.2	26.3	21.2	31.0	24.6	110.2
SD	2.25	5.90	9.06	7.89	7.73	19.4
最高	10	37.5	42.5	44.5	39	156
最低	0	3.5	0	0	0	38

表 1 1 8 8 年度問題間相関と採点者間一致度

	漢字	I	II	III	総点
漢字	1.00				
I	0.13	(0.47)			
II	0.08	0.18	(0.84)		
III	0.14	0.03	0.19	(0.91)	
総点	0.27	0.30	0.70	0.80	1.00
満点	10	40	60	90	200
平均	7.0	23.3	19.6	47.8	97.7
SD	2.08	3.12	10.10	12.80	18.85
最高	10	40	60	90	148
最低	0	3.5	0	0	34.5

表 1 2 8 9 年度問題間相関と採点者間一致度 (調整前)

	漢字	I	II	III	総点
漢字	1.00				
I	0.10	(0.74)			
II	0.10	0.31	(0.71)		
III	0.05	0.24	0.29	(0.84)	
総点	0.23	0.62	0.65	0.83	1.00
満点	10	60	50	80	200
平均	7.6	38.6	33.3	44.5	124.1
SD	2.16	5.85	5.79	10.71	16.80
最高	10	51	45.5	67.5	156.5
最低	0	0	11.5	8	34

表 1 3 8 9 年度問題間相関と採点者間一致度 (調整後)

	漢字	I	II	III	総点
漢字	1.00				
I	0.11	(0.85)			
II	0.10	0.31	(0.85)		
III	0.05	0.23	0.31	(0.92)	
総点	0.23	0.61	0.66	0.83	1.00
満点	10	60	50	80	200
平均	7.6	38.6	33.2	44.3	123.7
SD	2.16	5.91	5.93	10.86	17.05
最高	10	51	45.5	68	156
最低	0	0	10	8	34

表 1 4 9 0 年度問題間相関と採点者間一致度 (調整前)

	漢字	I	II	III	総点
漢字	1.00				
I	0.11	(0.60)			
II	0.03	0.10	(0.75)		
III	0.16	0.19	-0.06	(0.63)	
総点	0.29	0.60	0.55	0.68	1.00
満点	10	40	75	75	200
平均	6.9	14.4	31.5	40.7	93.5
SD	2.02	6.07	8.10	9.20	14.94
最高	10	32	56	60.5	140.5
最低	2	1.5	0	0	42.5

表 1 5 9 0 年度問題間相関と採点者間一致度 (調整後)

	漢字	I	II	III	総点
漢字	1.00				
I	0.13	(0.85)			
II	0.03	0.13	(0.84)		
III	0.15	0.19	-0.05	(0.78)	
総点	0.29	0.62	0.56	0.68	1.00
満点	10	40	75	75	200
平均	6.9	14.3	31.4	39.9	92.5
SD	2.02	6.35	8.09	9.24	15.29
最高	10	33.5	54.5	60	140.5
最低	2	1.5	0	0	41.5

となるであろう。そして、このテストの得点が全体としては、全く信頼性をもたないことは明らかであろう。

表1～15、特に表9以降は、問題間の相関がかなり低いことを示している。こうして、「論述」の信頼性は、採点者間の一致という観点からは高いが、問題間の相関という観点からは必ずしも高くないことになる。

仮に個々の問題が先の「コイン投げ」のテストとは異なり、それ自体としての信頼性は高かったとしても、問題間の相関が低いことは、テストとしては余り好ましいこととは言えない。このように相互相関の低い問題が無数に存在しうるとすれば、結局その中からどの問題が選ばれたか、ということによって、受験者の得点、ひいては相対順位に大きな変動が生ずることになるからである。しかし、この問題間相関が低さは、時間制限のある、ある程度大問主義をとらざるを得ない試験の宿命であると思われる。実際、数学や国語の場合もどうやら同じようなものであるらしい。

また、問題間相関が、採点者間の調整の後にも、ほとんど上昇していない点にも注意されるべきである。

しかし、よく見るとこの問題間相関にも、年度によってかなりの変動があることがわかる。特に、表14、15に示された90年度の結果は、見掛け上ほとんど上記の「コイン投げ」のテストにおいて予想される結果と異なるところがない。次に見る他の入試科目等との相関の低下とあわせて、やや気掛かりな点である。この点は、この年度から出題委員の人数を6名に増加し、各問題の出題者が完全に独立になったことも原因かもしれない。

(ただし、同じ体制で実施された91年度の結果(調整前)では、この問題間相関は88年度と89年度の中間の水準にまで回復している。)

入学試験における他の科目および高校調査書との関係

一般に、「測定したいと考えているもの」を測定している程度、という意味での妥当性を直接検討する方法はなく、当該のテストの得点と、他の情報との関連から推測していくことになる。信頼性の検討が、テストの内部構造の分析であるのに対し、妥当性の検討は、テストの外部構造の分析である。そこで、他の入試科目や、入学後の成績との関係を検討することには意味がある。

まず、論述的学力検査と他の科目との相関を、表16に示した。ここで、「調査書」は、高校における最終学年の全科目成績の平均値を（高校格差等を考慮せず）、本学の全受験者中の相対的位置（88年までは偏差値、89年以降は順位）に直し、改めて10段階の得点に変換したものである。

一般に、この種の相関係数は高いほど良いとは言えない。1つの科目として独立した試験を行う以上、他の科目では測定できない側面を反映していないと意味がないからである。しかし、論述的学力検査が、先に述べた「コイン投げ」のテストのようなものであったとすれば、どの科目との相関もほぼ0となってしまうであろう。それはもちろん好ましいこととは言えないから、この相関は低いほど良いとも言えない。

結局、この試験が果たすべき役割、ねらいから考えて、相関関係が適当なパターンを示しているかどうか、検討されるべきであろう。これらの相関の大部分は、少なくとも0とは高度に有意に異なっており、論述的学力検査の信頼性は皆無である、という最悪の状態が成立しているわけではないことは示されている。それ以上の問題として、論述的学力検査の本来のねらいから考えれば、国語、社会（1次）との相関が高いことが期待されるが、その点は必ずしもはっきりしない。ただし、この相関は、相手の科目の信頼性、受験者層の変動等、さまざまな要因の影響を受けるから、単純な解釈はできない。

なお、90年度に関しては、相関が全体として低下する傾向が明確に認められる。この点は、先に述べた問題間相関の低下と考えあわせると、今後の追跡調査において、慎重に検討すべき問題である。

次に、論述を含む各科目の、入試としての最終的な総得点との相関を表17に示した。これで見ると、論述的学力検査の総得点への寄与は、86年度と89年度を除き、全科目の最低である。87年度以降、理科、社会は各1科目、100点満点となったが、それにもかかわらず、これらの寄与は必ずしも低下していない（他の科目は、すべて200点満点である）。すなわち、論述的学力検査は、受験者の合否に対する規定力の最も低い科目である、と言ってよいであろう。

これが、単なる信頼性の（他の科目との比較における）相対的低さにすぎないのか、あるいは、「論述」独自の何らかの側面の測定に成功しているのか（つまり、何らかの妥当

表16 論述的学力検査と他の入試科目等との相関係数

科目	85年	86年	87年	88年	89年	90年
入試総点	0.51	0.56	0.53	0.51	0.55	0.42
1次総点	0.47	0.46	0.39	0.41	0.45	0.30
2次総点	0.47	0.55	0.60	0.53	0.56	0.46
調査書*	0.22	0.33	0.16	0.27	0.24	0.10
マルA	0.14	0.12	0.12	0.07	0.17	-0.09
1次						
国語	0.28	0.20	0.31	0.37	0.38	0.31
英語	0.30	0.54	0.38	0.30	0.36	0.23
理科	0.33	0.19	0.19	0.28	0.23	0.17
社会	0.42	0.39	0.31	0.37	0.30	0.25
数学	0.36	0.32	0.21	0.24	0.25	0.06
2次						
論述	—	—	—	—	—	—
英語	0.29	0.36	0.40	0.29	0.38	0.20
数学	0.27	0.22	0.16	0.16	0.18	0.10
性別	0.23	0.28	0.18	0.07	0.26	0.14

*最終学年全科目平均

表17 入試総点と各入試科目等との相関係数

科目	85年	86年	87年	88年	89年	90年
入試総点	—	—	—	—	—	—
1次総点	0.93	0.92	0.93	0.92	0.92	0.90
2次総点	0.90	0.88	0.90	0.89	0.91	0.89
調査書*	0.57	0.61	0.49	0.58	0.46	0.42
マルA	0.19	0.25	0.23	0.22	0.14	0.13
1次						
国語	0.55	0.50	0.61	0.66	0.52	0.49
英語	0.70	0.74	0.78	0.74	0.70	0.75
理科	0.69	0.64	0.63	0.62	0.53	0.56
社会	0.63	0.69	0.63	0.63	0.63	0.55
数学	0.72	0.67	0.70	0.70	0.71	0.60
2次						
論述	0.51	0.56	0.53	0.51	0.55	0.42
英語	0.70	0.65	0.78	0.70	0.75	0.76
数学	0.75	0.72	0.62	0.67	0.69	0.67
性別	0.16	0.18	0.26	0.28	0.05	0.22

*最終学年全科目平均

性があるのか), という点を明らかにするためには, さらに入学後の成績との関連を検討する必要がある.

入学後の成績との関係 1——教養部における成績

入学試験の妥当性の基準として, 最も適切なものは入学後の成績であると考えられている. これに関しては, まず, 教養部の成績との相関を, 表 18~22 に示した. なお, 成績は優を 3, 良を 2, 可を 1 として, 各カテゴリーごとに平均したものである (不可は, 本学では公式記録に残されない). これによれば, 論述的学力検査と教養部の成績との相関は, 85~87 年度入学者に関して言えば, 他の 1 次, 2 次諸科目より, 概して高い傾向にあった (修正前の欄). しかしながら, 前述のように, 論述的学力検査の合否への寄与は小さいから, この結果は選抜効果の影響を受けにくいことが理由であるとも考えられる. そこで, 相関係数を次の式で修正した.

$$r_{yz.x} = \frac{r_{yz} + r_{xy}r_{xz}(k-1)}{\sqrt{\{1+r_{xy}^2(k-1)\}\{1+r_{xz}^2(k-1)\}}}$$

ここで, X は選抜が行われる入試総点, k は, 選抜前後の X の分散の比を示す. この修正を行っても, 論述的学力検査の優位はほぼ動かない.

しかし, 88 年度に関しては, その予測力がかなり低下した. これが一過性のものかどうか, 今後の動向に注意する必要がある.

入学後の成績との関係 1——専門課程における成績

名古屋大学教育学部では, 1 年次から 2 年次にかけて, 五つの科目を全学生の必修として課している (教育学 I, II, 教育心理学概論 I, II, 教育統計学概説). これらの共通科目の成績, 専門課程における他の全教科の成績 (教職等に関する随意科目を除く), および, 卒業論文の成績を, 教養部の成績と同様の方法で数量化した.

また, 公式の成績以外の教官の主観的印象も重視すべきであると考えて, 学部の全教官 (助教授以上) に, 89~90 年度の卒業生 (大部分が, 85~86 年の入学者に相当する) のうち, 良く知っている学生について,

1. 総合的に見て優秀な学生だった.
2. 勉学に関する意欲が低かった.

3. 社会に出て力を発揮しそうである.
4. 真面目に努力した.
5. 基礎学力に欠ける面があった.
6. 研究者として可能性があった.

という6つの側面から、5段階で評定を求めた。その結果、89年度の卒業生56名については、平均3.6人(0-11人)、平成2年度の卒業生60名については、平均4.5人(1-10人)の教官による評価が得られた。その単純平均をその学生の評価値とした。

共通科目、専門科目、卒業論文と各入試科目等との相関係数、および、上の主観的評価の相関係数を、表23~31に示した。この表の値は、教養部の成績との相関と比べて、全体的に低いものとなっている。しかし、85年入学者については、全般に予想される方向の結果であると言えよう。

しかし、86年度入学者については、入学試験全体として、入学後の成績の予測力は極めて低く、若干の逆相関すら認められる。論述的学力検査は、「専門全体」の成績については、なお、受験科目中では高い相関を示している。実際、86年度は教養部の成績の時点から、入試成績との相関は極めて低かったから、その中ではむしろ高い予測力を示していると言ふべきかもしれない。

こうした、特定の年度における相関の低下が生ずる理由はあきらかでない。クラスの雰囲気といった、学生間の相互作用が何らかの影響をもっているのかもしれない。

専門課程における成績との関係についての資料は、現在のところ2年分しか得られておらず、一定の結論を得るためには、今後なお資料の蓄積が必要である。

5. おわりに

小規模学部の入学試験方法の検討であるため、どうしてもサンプルが小さくならざるを得ず、そのために標本誤差が大きいこともあって、明確な結論を得ることは難しい。今のところ言えることは、論述的学力検査は、他の科目とは異なった能力の側面を測定しており(つまり、単なる誤差によって、他の科目との相関が低いわけではなく)、それは必ずしも見当はずれのものでもなさそうだ(入学後の成績と無関係ではない)といったところである。論述的学力検査のねらい、出題傾向も変化しつつあり、また、入試制度全体の変

表18 教養部人文科学の成績との相関

科目	85年度		86年度		87年度		88年度	
	修正前	修正後	修正前	修正後	修正前	修正後	修正前	修正後
入試総点	0.35	0.74	0.04	0.09	0.24	0.57	0.22	0.63
1次総点	0.20	0.69	-0.08	0.02	0.15	0.50	0.40	0.68
2次総点	0.32	0.73	0.13	0.14	0.16	0.54	-0.12	0.51
調査書*	0.39	0.69	0.25	0.25	0.34	0.44	0.33	0.46
マルA	0.39	0.64	0.16	0.17	0.27	0.31	0.27	0.42
1次								
国語	-0.01	0.18	0.03	0.06	-0.12	0.14	0.36	0.57
英語	-0.11	0.06	-0.08	-0.03	0.26	0.50	0.14	0.36
理科	0.26	0.62	-0.07	0.01	0.20	0.46	0.19	0.49
社会	0.11	0.24	-0.03	0.02	-0.14	-0.03	0.04	0.48
数学	0.08	0.58	-0.06	0.00	0.20	0.46	0.13	0.42
2次								
論述	0.37	0.66	0.28	0.28	0.30	0.55	-0.12	0.34
英語	0.01	0.40	0.03	0.06	0.11	0.47	0.10	0.53
数学	0.22	0.67	0.04	0.09	0.48	0.34	-0.15	0.40
性別	0.48	0.67	0.57	0.56	0.37	0.22	0.55	0.61

*最終学年全科目平均

表19 教養部社会科学の成績との相関

科目	85年度		86年度		87年度		88年度	
	修正前	修正後	修正前	修正後	修正前	修正後	修正前	修正後
入試総点	0.25	0.60	0.08	0.17	0.05	0.15	0.08	0.27
1次総点	0.23	0.59	-0.18	0.01	-0.02	0.10	0.27	0.36
2次総点	0.15	0.57	0.29	0.28	0.07	0.16	-0.16	0.17
調査書*	0.44	0.66	0.23	0.27	0.09	0.13	0.23	0.30
マルA	0.47	0.65	0.16	0.19	0.15	0.17	0.17	0.25
1次								
国語	-0.15	0.03	0.03	0.09	0.04	0.10	0.18	0.28
英語	0.02	0.13	-0.09	0.00	0.10	0.16	0.05	0.15
理科	0.28	0.56	-0.22	-0.06	0.01	0.09	0.11	0.25
社会	0.14	0.24	-0.19	-0.08	-0.08	-0.05	0.21	0.32
数学	0.07	0.48	-0.01	0.08	-0.11	0.00	0.12	0.24
2次								
論述	0.33	0.58	0.26	0.30	0.32	0.32	0.07	0.23
英語	0.11	0.40	0.09	0.14	-0.02	0.10	0.00	0.20
数学	0.00	0.48	0.19	0.23	-0.09	0.04	-0.22	0.07
性別	0.23	0.47	0.59	0.58	0.26	0.23	0.33	0.38

*最終学年全科目平均

表 2 0 教養部自然科学の成績との相関

科目	85年度		86年度		87年度		88年度	
	修正前	修正後	修正前	修正後	修正前	修正後	修正前	修正後
入試総点	0.23	0.58	0.05	0.10	0.31	0.69	0.05	0.18
1次総点	0.11	0.52	-0.01	0.06	0.11	0.56	0.22	0.27
2次総点	0.23	0.58	0.08	0.12	0.27	0.67	-0.14	0.10
調査書*	0.34	0.59	0.17	0.19	0.34	0.45	0.13	0.19
マルA	0.46	0.64	0.17	0.18	0.22	0.27	0.18	0.23
1次								
国語	-0.14	0.04	0.11	0.14	-0.09	0.22	0.24	0.30
英語	0.08	0.17	-0.21	-0.13	0.11	0.46	-0.19	-0.10
理科	0.18	0.49	0.14	0.16	0.26	0.55	0.12	0.20
社会	0.20	0.29	0.04	0.08	0.10	0.17	0.04	0.16
数学	-0.09	0.38	-0.16	-0.08	0.00	0.40	0.18	0.24
2次								
論述	0.16	0.46	0.14	0.16	0.15	0.53	0.00	0.11
英語	0.08	0.37	-0.13	-0.08	0.22	0.60	-0.13	0.06
数学	0.14	0.52	0.09	0.13	0.09	0.50	-0.08	0.09
性別	0.23	0.46	0.26	0.26	0.05	-0.06	0.49	0.52

*最終学年全科目平均

表 2 1 教養部英語の成績との相関

科目	85年度		86年度		87年度		88年度	
	修正前	修正後	修正前	修正後	修正前	修正後	修正前	修正後
入試総点	0.21	0.54	0.15	0.31	0.13	0.35	0.14	0.45
1次総点	-0.01	0.44	0.06	0.25	0.02	0.27	0.00	0.36
2次総点	0.30	0.56	0.16	0.31	0.13	0.35	0.13	0.44
調査書*	0.30	0.54	0.31	0.39	0.19	0.27	0.31	0.42
マルA	0.29	0.50	0.31	0.36	0.13	0.17	0.20	0.32
1次								
国語	-0.10	0.06	0.05	0.17	-0.24	-0.06	0.04	0.27
英語	0.20	0.27	0.03	0.17	0.19	0.34	0.30	0.43
理科	0.15	0.45	0.09	0.25	0.16	0.32	-0.19	0.13
社会	-0.05	0.07	-0.03	0.13	-0.08	-0.02	0.01	0.33
数学	-0.14	0.32	0.00	0.15	0.13	0.30	-0.12	0.15
2次								
論述	0.18	0.45	0.22	0.32	0.09	0.29	0.05	0.32
英語	0.16	0.40	0.25	0.33	0.12	0.33	0.25	0.48
数学	0.16	0.50	0.01	0.21	0.02	0.24	-0.03	0.32
性別	0.42	0.58	0.39	0.36	0.16	0.09	0.16	0.28

*最終学年全科目平均

表 2 2 教養部第 2 外国語の成績との相関

科目	85年度		86年度		87年度		88年度	
	修正前	修正後	修正前	修正後	修正前	修正後	修正前	修正後
入試総点	0.40	0.79	0.02	0.04	0.20	0.50	0.20	0.60
1次総点	0.23	0.73	-0.09	-0.03	0.21	0.49	0.00	0.47
2次総点	0.36	0.78	0.11	0.09	0.06	0.44	0.20	0.59
調査書*	0.52	0.77	0.43	0.40	0.27	0.37	0.26	0.41
マルA	0.36	0.64	0.44	0.43	0.10	0.16	0.12	0.30
1次								
国語	0.05	0.23	-0.08	-0.05	-0.11	0.12	0.01	0.32
英語	0.23	0.30	0.11	0.11	0.18	0.41	0.19	0.38
理科	0.12	0.58	-0.01	0.01	0.20	0.43	-0.07	0.31
社会	-0.08	0.12	-0.14	-0.10	0.14	0.19	0.05	0.46
数学	0.10	0.62	-0.08	-0.05	0.17	0.40	-0.11	0.24
2次								
論述	0.22	0.60	0.14	0.13	0.21	0.46	-0.09	0.33
英語	0.25	0.56	0.11	0.11	0.13	0.43	0.21	0.56
数学	0.16	0.69	0.03	0.04	-0.14	0.24	0.14	0.52
性別	0.42	0.64	0.48	0.48	0.24	0.13	0.42	0.51

*最終学年全科目平均

表 2 3 専門課程における成績との相関 (1)

科目	共通科目			
	85年度		86年度	
	修正前	修正後	修正前	修正後
入試総点	0.21	0.54	0.04	0.09
1次総点	0.09	0.49	-0.15	-0.03
2次総点	0.21	0.54	0.20	0.18
調査書*	0.38	0.59	0.26	0.27
マルA	0.38	0.56	0.28	0.29
1次				
国語	0.05	0.23	0.15	0.16
英語	0.15	0.27	-0.13	-0.06
理科	0.21	0.49	-0.13	-0.05
社会	-0.02	0.10	-0.08	-0.03
数学	-0.12	0.31	-0.19	-0.11
2次				
論述	0.07	0.38	0.22	0.23
英語	0.28	0.51	0.10	0.12
数学	0.03	0.43	0.11	0.13
性別	0.16	0.35	0.64	0.63

*最終学年全科目平均

表 2 4 専門課程における成績との相関 (2)

科目	専門全体**			
	85年度		86年度	
	修正前	修正後	修正前	修正後
入試総点	0.26	0.62	0.03	0.06
1次総点	0.11	0.56	0.06	0.08
2次総点	0.27	0.62	-0.01	0.04
調査書*	0.37	0.62	0.40	0.38
マルA	0.32	0.55	0.30	0.31
1次				
国語	-0.20	0.08	-0.02	0.01
英語	0.30	0.40	0.03	0.06
理科	0.18	0.53	0.00	0.04
社会	0.08	0.20	0.07	0.09
数学	-0.10	0.39	0.09	0.10
2次				
論述	0.25	0.54	0.32	0.30
英語	0.22	0.51	0.07	0.09
数学	0.07	0.52	-0.14	-0.05
性別	0.24	0.44	0.17	0.17

*最終学年全科目平均

** 共通科目除く

表 2 5 専門課程における成績との相関 (3)

科目	卒業論文			
	85年度		86年度	
	修正前	修正後	修正前	修正後
入試総点	0.23	0.58	-0.10	-0.20
1次総点	0.28	0.59	-0.10	-0.20
2次総点	0.09	0.53	-0.06	-0.18
調査書*	0.22	0.51	0.17	0.06
マルA	0.20	0.45	-0.04	-0.08
1次				
国語	-0.04	0.18	-0.25	-0.30
英語	0.19	0.31	0.01	-0.09
理科	0.29	0.56	-0.08	-0.18
社会	0.04	0.16	-0.08	-0.17
数学	0.04	0.43	0.12	0.00
2次				
論述	-0.07	0.31	0.03	-0.06
英語	0.22	0.49	-0.07	-0.15
数学	-0.03	0.44	-0.04	-0.16
性別	0.29	0.46	-0.15	-0.14

*最終学年全科目平均

表26 教官の評定との相関(1)

科目	評定(総合的に優秀)			
	85年度		86年度	
	修正前	修正後	修正前	修正後
入試総点	0.25	0.61	0.00	0.00
1次総点	0.14	0.56	0.00	0.00
2次総点	0.22	0.59	0.00	0.00
調査書*	0.19	0.50	0.24	0.21
マルA	0.07	0.38	0.16	0.16
1次				
国語	-0.02	0.20	-0.28	-0.25
英語	0.10	0.24	-0.01	-0.01
理科	0.26	0.56	-0.02	-0.01
社会	0.10	0.21	-0.02	-0.02
数学	-0.10	0.37	0.26	0.22
2次				
論述	0.09	0.43	0.17	0.15
英語	0.23	0.51	-0.03	-0.02
数学	0.06	0.50	-0.04	-0.03
性別	0.32	0.49	-0.06	-0.06

*最終学年全科目平均

表27 教官の評定との相関(2)

科目	評定(勉学への意欲低)			
	85年度		86年度	
	修正前	修正後	修正前	修正後
入試総点	-0.29	-0.67	0.24	0.45
1次総点	-0.28	-0.66	0.23	0.44
2次総点	-0.17	-0.63	0.13	0.39
調査書*	-0.35	-0.63	-0.07	0.13
マルA	-0.16	-0.47	0.16	0.24
1次				
国語	-0.07	-0.29	0.32	0.45
英語	-0.12	-0.27	0.05	0.26
理科	-0.26	-0.60	0.10	0.33
社会	-0.14	-0.25	0.19	0.37
数学	0.02	-0.47	-0.02	0.20
2次				
論述	-0.14	-0.49	0.05	0.23
英語	-0.29	-0.58	0.08	0.25
数学	0.04	-0.51	0.10	0.36
性別	-0.38	-0.55	0.04	0.00

*最終学年全科目平均

表 2 8 教官の評定との相関 (3)

科目	評定 (社会に出て力発揮)			
	85年度		86年度	
	修正前	修正後	修正前	修正後
入試総点	0.22	0.56	-0.12	-0.24
1次総点	0.16	0.53	-0.17	-0.27
2次総点	0.17	0.53	-0.01	-0.17
調査書*	0.10	0.42	0.10	-0.02
マルA	0.10	0.37	-0.02	-0.07
1次				
国語	-0.08	0.14	-0.30	-0.36
英語	0.10	0.24	-0.17	-0.26
理科	0.26	0.53	-0.21	-0.29
社会	0.04	0.15	0.03	-0.09
数学	-0.01	0.39	0.15	0.00
2次				
論述	-0.08	0.29	-0.04	-0.13
英語	0.07	0.38	0.01	-0.10
数学	0.15	0.51	0.00	-0.16
性別	0.10	0.31	-0.24	-0.22

*最終学年全科目平均

表 2 9 教官の評定との相関 (4)

科目	評定 (真面目に努力)			
	85年度		86年度	
	修正前	修正後	修正前	修正後
入試総点	0.25	0.60	-0.13	-0.26
1次総点	0.21	0.58	-0.16	-0.28
2次総点	0.16	0.57	-0.04	-0.21
調査書*	0.35	0.60	0.09	-0.04
マルA	0.15	0.43	0.05	-0.01
1次				
国語	0.06	0.26	-0.11	-0.20
英語	0.24	0.35	-0.08	-0.19
理科	0.18	0.52	-0.15	-0.26
社会	0.09	0.20	-0.20	-0.29
数学	-0.08	0.38	0.07	-0.07
2次				
論述	0.02	0.38	0.04	-0.08
英語	0.24	0.51	0.00	-0.11
数学	0.01	0.48	-0.06	-0.21
性別	0.41	0.56	0.14	0.15

*最終学年全科目平均

表 3 0 教官の評定との相関 (5)

科目	評定 (基礎学力欠く)			
	85年度		86年度	
	修正前	修正後	修正前	修正後
入試総点	-0.19	-0.50	0.00	0.01
1次総点	-0.28	-0.54	0.07	0.05
2次総点	-0.03	-0.44	-0.06	-0.03
調査書*	-0.25	-0.48	-0.34	-0.30
マルA	-0.16	-0.39	-0.28	-0.27
1次				
国語	-0.02	-0.20	0.13	0.12
英語	-0.07	-0.20	0.04	0.03
理科	-0.33	-0.54	0.06	0.05
社会	-0.32	-0.38	0.03	0.03
数学	0.12	-0.29	-0.05	-0.04
2次				
論述	-0.31	-0.52	-0.02	-0.02
英語	-0.20	-0.44	-0.07	-0.06
数学	0.19	-0.30	-0.03	-0.02
性別	-0.33	-0.47	-0.14	-0.14

*最終学年全科目平均

表 3 1 教官の評定との相関 (6)

科目	評定 (研究者の可能性)			
	85年度		86年度	
	修正前	修正後	修正前	修正後
入試総点	0.22	0.54	-0.18	-0.36
1次総点	0.16	0.52	-0.21	-0.37
2次総点	0.16	0.52	-0.07	-0.29
調査書*	0.21	0.48	0.31	0.11
マルA	0.05	0.33	0.11	0.02
1次				
国語	-0.16	0.07	-0.21	-0.33
英語	0.24	0.35	-0.07	-0.23
理科	0.25	0.52	-0.22	-0.36
社会	0.18	0.27	-0.10	-0.25
数学	-0.14	0.31	0.05	-0.14
2次				
論述	0.02	0.35	0.10	-0.07
英語	0.22	0.47	-0.14	-0.26
数学	0.03	0.44	-0.06	-0.27
性別	0.06	0.28	-0.20	-0.17

*最終学年全科目平均

化もあって、事態は流動的であり、その中で論述的学力検査の妥当性を明らかにするためには、今後とも検討を継続する必要がある。

なお、試験の「質」には必ずしも相関関係だけで云々できない面がある。例えば、ある科目を課すことにより、受験者の母集団が変化する可能性がある。また、その科目のための準備が、大学教育に対する適切な準備活動として機能する可能性もある（こういったテストは、本質的妥当性が高いと言えよう）。こうした点がもし事実であるとすれば、いずれにしても、受験者間の差は縮小するから、相関係数にもとづく、信頼性と妥当性は、一見低下することになる。これらの点については、今のところ確認することが難しい。

そういった点も総合的に考慮した上で、あえて現時点での評価を下すとすれば、この論述的学力検査は、数多くの小論文的な入学試験科目の中では、かなりの成功をおさめたものと言えるのではなかろうか。その理由としては、最初にも述べたような、あくまでも学力検査として、その「質」の確保に努力してきたことがあげられよう。これは課題型の小論文と比較して、簡単な対策が立てにくく、その結果一定の妥当性が維持されてきたと考えられる。また一方、高校卒業者を対象とする試験として、余りにもユニークになりすぎることを回避してきたことも、成功の理由の一つにあげることができよう。他大学における資料型小論文とある程度共通の準備が可能であることは、受験母集団の極端なローカル化を回避することに役立った可能性が高い。

最後に、この論述的学力検査については、比較的小規模なサンプルとは言え、かなり多面的で信頼し得るデータを一定期間にわたって蓄積できた（さらに蓄積されつつある）ことは、将来の入試研究と入試改革に対し、一定の寄与となりうるであろう。