

# マークシート形式と記述形式による数学の「学力」

村上 隆, 三宅正武, 藤村宣之, 浪川幸彦, 鈴木浩志, 鈴木紀明 (名古屋大学)  
田栗正章, 内田照久 (独立行政法人大学入試センター)

大学入試センター試験で用いられているようなマークシート形式の試験が、果たして受験生の数学の能力を測定するのに十分なものであるかどうかについては、従来から多くの疑問が呈されてきた。本研究は、同一の問題をマークシート形式と記述形式で作成し、同一の受験者群に実施することにより、その問いに答えようとしたものである。同時に実施された受験者の学習観に関する意識調査の結果と関係づけることにより、2つの形式で測定される能力の違いについて、新たな知見が得られることが期待される。

## 1. 問題の所在

### 1.1 基本的な問い

大学入試センター試験の成績、あるいは一般にマークシートを用いた客観形式の試験は、受験者のどのような能力を反映しているのだろうか。それは、わが国の入学試験におけるより古い形態であり、現在も多くの大学の個別試験で使われている記述形式の試験とは、別の能力を測定しているのだろうか。それとも、2つの形式の試験の差は、単なる難易度の差に過ぎないのだろうか。数学の入学試験に領域を限定した上で、この問題に対してある程度の解答を与えることが、本研究の目的である。

### 1.2 マークシート形式への疑問

いろいろと制約の多いマークシート形式の試験であるが、大学入試センター試験では、単なる記憶力以上のものを評価するために多大な努力が払われている。しかし、一定の長さの論理的推論過程を必要とする数学の試験では、その論証過程の筋道が問題上に示され、事実上その空白を埋めていくに過ぎないマークシート形式では、計算力以上の論理的思考力を問うことはできないという批判が根強くある。実際、この批判に反論することは難しく、マークシート形式を擁護する人たちも、たとえば「大学

入試センター試験は、個別試験と組み合わせることによって、選抜の指標として機能する」と主張することが多く、マークシート形式の試験だけで十分であるという論をなす人は少ない。

### 1.3 マークシート形式で測られる能力

少なくとも、大学入試センター試験では、記述形式の試験で測定できる数学の能力の一部を測定できないことは、認めざるを得ないであろう。具体的に言えば、マークシート形式の試験では、計算能力と（主として公式に関する）知識を測ることはできるが、論理的思考能力は測れないということである。その一方で、多くの数学教育関係者が、大学入試センター試験の数学の問題が、よく練り上げられた優れた問題であることに合意しているのも確かである。つまり、せっかく時間をかけて丁寧に作られた良問が、マークシートによる解答という形式をとることによってその優れた性質の多くを失ってしまっていることになる。もしそうであるとすれば、現行の大学入試センター試験の問題を、記述形式にもどして実施すれば、いわゆるレベルの高い大学においても適切な入試問題として使える可能性がある。

このように、大学入試センター試験が個

別試験と比べて何かを欠いているという指摘がある一方で、大学入試センター試験の「数学」が、高等教育の基礎となるには不必要な、スピードと正確さに関する過剰なトレーニングを含み、少なくともそれが競争試験として用いられる限り、むしろ有害であるという指摘もある。その意味では、大学入試センター試験の「数学」は、記述式試験に対して欠落している部分だけでなく、余剰の部分をも含んでいるということになる。

現在までのところ、マークシート形式の試験と記述形式の試験とが、どのような意味で相補い合うものとなるか、あるいは、マークシート形式の試験は能力評価に不適切な面を含んでいるのかについては、かならずしも明確な答えが示されているわけではない。そのためには、より踏み込んだ議論と、何よりもその基礎資料となる実証的なデータが必要とされよう。しかしながら、この問題についての実証的なデータにもとづく議論は極めて少ない。次にそのような少数の例外について検討してみよう。

## 2 先行する諸研究とその結果

### 2.1 森田の指摘

森田 (2003) は、東北大学におけるセンター試験の結果について、次のような現象を報告した。すなわち、「数学」の大学入試センター試験の得点を横軸に、同じく個別試験の得点を縦軸にとってプロットすると、次のような特徴が見出される。

- (1) センター試験において、満点に近い高得点者が多く、著しく偏った分布となっている。
- (2) センター試験、個別試験の0点と満点を結ぶ直線の右下に分布が偏っている。すなわち、受験者の個別試験の得点は、センター試験の得点の水準を越えることはほとんどない。

(以上の表現は、西田 (2003) にもとづく。)

この現象について森田は、「このようになった理由は不明」であるとしながらも、論理力や思考力が重視される「数学」の試験では、答えだけを書くマークシートによる解答には「偶然性が入り、双方の間の相関が下がっている」という解釈が成り立つかもしれないと論じている。

西田 (2003) は、自身の所属する大学 (京都大学) においても、「数学」の散布図に年度、学部を問わず同様の現象が見られることを確認した上で、上記、(1) と (2) にわけて、それぞれに対する解釈を述べている。すなわち、(1) については、(京都大学や東北大学のようなレベルの大学において、) 大学入試センター試験が易しすぎることを意味しているとする。他方、(2) については、森田の解釈、すなわち「センター試験が・・・数学の能力の一部だけを問うているからなのか、あるいは他の要素があるからなのかは即断はできない」とし、「2次試験の難度、あるいは試験問題がどのような能力を問うているかを考慮に入れ、何年かのデータを解析する必要がある」としている。

現象の解釈については、改めて検討するとして、名古屋大学における分析の結果を見ても、個別試験の問題が難しかったときには、外国語においても (2) の特徴をもつ分布の形が現れることがわかっており、この現象が問題の難易度と切り離して論じられないことは確かである。

### 2.2 試験問題の記述形式への書き換え

ここで、マークシート形式と記述形式の違いを明らかにするための、ある意味でより厳密な方法として、センター試験の問題を記述形式に書き換えて行われた研究について述べておこう。筆者の知る限り、そうした研究で公表されているものは、桜井・藤井・岩坪・伊藤・松田 (2003) が唯一である。彼らは、「数学」ではなく「物理」

の問題を記述式に書き換え、それを普通科と工業科の高校生合わせて 119 名に実施した。結果はやはり、前記 (2) と同様の傾向を示している。

### 3 本研究の目的と調査の計画

#### 3.1 何が知りたいのか

このような過去の研究成果の上に立って、マークシート形式と記述形式の試験を比較したり、その間の関係を検討したりする際には、どのような観点から、何を指標としたらよいのかを考えてみよう。主な観点は、次の 2 つであると思われる。

- (1) マークシート形式と記述形式の難易度の差
- (2) マークシート形式と記述形式の得点間の相関関係

それぞれの観点について、若干説明しておきたい。まず、(1) については、一般にマークシート形式の試験の方が、難易度が低いことは想像できるものの、どのような設問において、どの程度の得点上昇が得られるかという点についてのデータにもとづく実証的情報は従来ほとんど得られていない。したがって、マークシート形式の問題と、それを書き換えた記述形式の問題を、同程度の能力をもつと考えられる 2 つの受験者のグループに対して実施し、成績を比較することには十分に意味がある。

しかしながら、センター試験と個別試験の測定内容について検討する観点からは、(2) の方がより重要である。すなわち、多数の受験者からマークシート形式の問題に対する得点と、記述形式の問題に対する得点の両方が得られれば、その結果にもとづいて描かれる散布図から、多くの情報が得られる可能性がある。

#### 3.2 散布図の検討

まず、森田らが見出したような、散布図上の右下三角部分に点が集中するような現

象が見られるかどうかの問題になる。この形態は全受験者の結果には見られるが、合格者だけの結果では原点付近の受験者が取り除かれるために、完全な形では見られなくなる。大学生を対象とする場合、受験者は合格者の中から募集されるので、必ずしも明確な下三角の分布形態は得られない可能性がある。

さらに、ある程度こうした結果が得られたとしても、その解釈は必ずしも簡単ではない。たとえば、次のような複数可能性を考えることができる。

まず、森田の指摘のように、大学入試センター試験で問われているのは、数学の能力の一部でしかないことを示すという解釈がある。個別試験では、マークシート形式では問えない能力が試される結果、それを欠いている受験生は、散布図の右下方向、すなわち、センター試験の得点は高くても個別試験の得点が低い領域に分布することになる。逆に、センター試験でしか測りえない能力というのはないので、左上部分には点が存在しない。

やや異なった観点から、筆者はこの現象について、次のように考えたことがある。すなわち、大学入試センター試験の能力は個別試験の必要条件である。したがって、個別試験のトレーニングをすれば、センター試験の得点はそのレベルまで上がる。つまり、センター試験だけのためのトレーニングは無駄であり、少なくとも受験準備に関する限り、個別試験への対応に努力を集中した方がよい。

他方で、この傾向を全く 2 つの試験の難易度の差によって説明しようとする考え方もある。すなわち、この種のデータが集められた大学では、大学入試センター試験は易しすぎ、天井効果 (ceiling effect) を引き起こす。そのため、個別試験との関係が非線形的になった結果であるとするもの

である。最近では、この現象が必ずしもいわゆる理系科目に限らず、外国語等においても起こることが発見されていることから、この現象が、必ずしも教科科目の特性によらない可能性は示唆されている。

### 3.3 測定内容の差を明らかにするために

数学のマークシート形式の試験と記述形式の試験は異なるものを測っているのか、

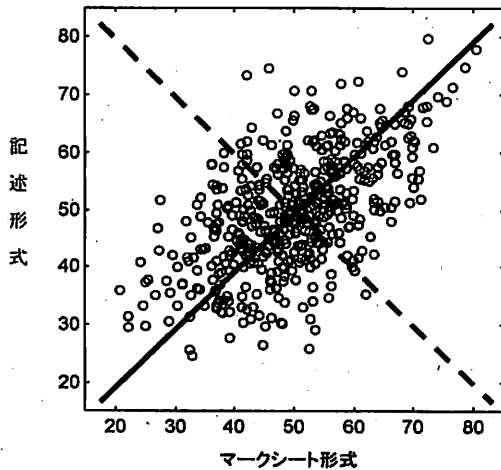


図1 マークシート形式(横軸)と記述形式(縦軸)の得点間に想定される関係。

2つの得点はともに平均値が50、標準偏差が10の正規分布で、相互相関がほぼ0.6となるように生成されている。

あるいは、内容的には同じで難易度だけが異なるのかというのが、われわれが本研究で明らかにしたいことである。2つの形式のテストの間の散布図にみられる特徴的な分布形態は、この問題の解決への1つのヒントを与えているが、これだけで2つのテストの間の内容の違いを立証する証拠とはなりえない。マークシート形式と記述形式の内容的な違いを明らかにするには、2つのテストの得点以外の情報が必要である。

そこで、一旦2つの形式の試験の得点分布に関する特徴については白紙に戻し、両形式のテストから得られた得点の間に、通

常の2変量正規分布が成立するという状態を想定して、問題を整理してみよう。このような状態は、何らかの受験母集団に対し、適切な難易度のマークシート形式のテストと記述形式のテストを作成しえたという想定であるが、大学入試センター試験に課されたさまざまな制約条件を外すことのできる状況を考えれば、こうしたことは原理的に不可能というわけでもないであろう。その状態を示したのが図1である。

ここでまず、左下から右上に45度の角度で引かれた直線への、各点からの正射影を考えると、これは2つの得点の和に対応するものとなっており、いわば個人の総合力に当たるものと考えてよいであろう。他方、左上から右下への右下がりの破線に対する正射影を考えると、こちらは(右下方向を正と考えると)記述の得点からマークの得点を引いたものに当たる。すなわち、結果的に記述が得意か、マークが得意かを(あくまでも相対的に、かつ総合力とは無関係に)あらわしたものと見ることができる。

### 3.4 得点差の規定因

マークシート形式と記述形式の得点差が存在するとすれば、その規定因にはどのようなものが考えられるであろうか。実際の観測可能性も考えて、われわれは主に次の2つを検討することとした。

- (1) 記述試験における解答の内容。とくに、誤答のあり方。
- (2) 受験者の数学の学習観、あるいは数学の学習方略とでも言うべき意識と行動のあり方。

それぞれについて、若干の説明をしておこう。右下三角部分に集中する得点分布の特徴として特に注目されるのは、マークシート形式の得点は高い(満点、あるいはそれに近い)受験者の間で、記述形式の得点に大きな分散が存在することである。したがって、マークシートが高得点であるのに

記述の得点が低い受験者の、記述形式の部分の答案を検討することから、マークシートでは測りえない数学の能力が何であるかを具体的に明らかにできる可能性がある。また、彼らの学習観や学習方略の特徴を明らかにすることから、こうした受験者を減らしていくための教授法レベルでの方策が明らかになる可能性がある。

### 3.5 研究計画の概要

以上のような目的からすれば、だいたい次のような調査プログラムが浮かび上がる。

- (1) 大学入試センター試験に匹敵する内容と難易度をもつマークシート形式の試験と、それに対応する内容をもつ記述形式の試験を、同一の受験者に同一時期に実施する。
- (2) 数学の学習観や学習方略を測定するための質問紙を作成し、同じ受験者に同時期に実施する。

具体的には、2005年度大学入試センター試験追試験における、『数学Ⅰ・A』と『数学Ⅱ・B』の追試験問題を、記述形式に書き直し、それらと本来のマークシート形式の問題を、名古屋大学の理系学部1年生に実施した。また、数学の学習観と数学の学習方略に関する質問紙調査を作成し、同じ学生に実施した。

## 4 受験者の募集と調査の実施

**受験者** 受験者の募集は、2005年12月8日(木)、9日(金)の両日、理学部と工学部対象の名古屋大学理系基礎科目「数学」の授業において、授業担当教員を通じて行われた。実施を予定した教室の大きさ等から350名に達したところで募集を打ち切った。

**受験者の割り当て** 第2章で見たように、オリジナルのマークシート形式の問題と、書き換えられた記述形式の問題とを組み合

表1 受験番号と解答した問題の対照表

|       | マークシート形式                            | 記述形式                                |
|-------|-------------------------------------|-------------------------------------|
| 100番台 | IAQ2<br>IIBQ3 IIBQ4 IIBQ5           | IAQ1 IAQ3 IAQ4<br>IIBQ1 IIBQ2       |
| 200番台 | IAQ2<br>IIBQ1 IIBQ2                 | IAQ1 IAQ3 IAQ4<br>IIBQ3 IIBQ4 IIBQ5 |
| 300番台 | IAQ1 IAQ3 IAQ4<br>IIBQ3 IIBQ4 IIBQ5 | IAQ2<br>IIBQ1 IIBQ2                 |
| 400番台 | IAQ1 IAQ3 IAQ4<br>IIBQ1 IIBQ2       | IAQ2<br>IIBQ3 IIBQ4 IIBQ5           |

数Ⅰ・Aの第1問をIAQ1等と略記した

わせて、数Ⅰ・Aと数Ⅱ・Bのそれぞれについて2種の問題冊子が作られた。それぞれの問題冊子をIA-1, IA-2, IIB-1, IIB-2と呼ぶことにし、受験者を所属部局が偏らないように配慮して4分して、次のように割り当てた。

|              |              |
|--------------|--------------|
| 受験番号 101-187 | IA-1 と IIB-1 |
| 受験番号 201-287 | IA-1 と IIB-2 |
| 受験番号 301-387 | IA-2 と IIB-1 |
| 受験番号 401-486 | IA-2 と IIB-2 |

受験者総数が350に満たないのは、この作業を行う時点で既に辞退者があったからである。受験番号と解いた問題との関係を一覧表にしたものが表1である。このうち、IAQ1, IAQ2, IIBQ1, IIBQ2が必答、他は選択である。

**試験の実施** 試験は2006年1月15日、名古屋大学経済学部カンファレンスルームで実施された。試験当日には、さらに約30名が欠席したが、著しい遅刻等、目だった問題はなく、円滑に実施することができた。

試験時間等は、以下のようにした。

|             |             |
|-------------|-------------|
| 10:30-11:50 | 数Ⅰ・A (80分)  |
| 13:00-14:30 | 数Ⅱ・B (90分)  |
| 14:45-15:15 | 質問紙調査 (30分) |

マークシート部分の採点は、学内のマークシートリーダーを用いてExcelファイルに直した後、Excelの関数を用いて行わ

れた。記述形式の部分の採点は、名古屋大学多元数理研究科所属のOD, PDレベルの6名の採点者によって行われた。なお、記述形式の部分の答案は、今後の質的分析のために、すべてカラーの pdf ファイルとして保存された。

5 結果の概要

5.1 難易度等の比較

マークシート形式の得点と、記述形式の得点を、それぞれ、小問ごとに得点を集計したものの平均値と標準偏差を満点(FULL と表示)、受験者数

表2 マークシート形式と記述形式の小問ごとの比較

| 問題 | マークシート  |       |       |      | 記述   |       |       |      |     |
|----|---------|-------|-------|------|------|-------|-------|------|-----|
|    | FULL    | MEAN  | SD    | N    | FULL | MEAN  | SD    | N    |     |
| 1  | 40      | 34.59 | 5.88  |      | 40   | 30.61 | 7.83  |      |     |
|    | [1]     | 20    | 18.19 | 3.16 |      | 20    | 15.13 | 5.31 |     |
|    | [1] (1) | 15    | 14.47 | 1.49 | 160  | 12    | 9.88  | 3.21 | 158 |
|    | [1] (2) | 5     | 3.72  | 2.19 |      | 8     | 5.25  | 3.04 |     |
|    | [2]     | 20    | 16.40 | 4.38 |      | 20    | 15.47 | 5.78 |     |
| 2  | 40      | 33.22 | 7.11  |      | 40   | 24.46 | 8.27  |      |     |
|    | [1]     | 20    | 17.09 | 3.64 |      | 20    | 13.20 | 4.47 |     |
|    | [1] (1) | 12    | 11.11 | 2.29 | 158  | 12    | 8.74  | 3.79 | 160 |
|    | [1] (2) | 8     | 5.97  | 2.38 |      | 8     | 4.46  | 2.19 |     |
|    | [2]     | 20    | 16.13 | 5.35 |      | 20    | 11.26 | 5.96 |     |
| 3  | 20      | 14.67 | 4.98  |      | 20   | 11.49 | 7.11  |      |     |
|    | (1)     | 8     | 5.58  | 2.38 | 141  | 8     | 4.92  | 3.16 | 141 |
|    | (2)     | 12    | 9.09  | 3.79 |      | 12    | 6.57  | 5.21 |     |
| 4  | 20      | 15.05 | 4.74  | 22   | 20   | 10.32 | 7.11  | 22   |     |
| 1  | 30      | 24.29 | 4.82  |      | 30   | 11.51 | 8.75  |      |     |
|    | [1]     | 16    | 11.91 | 3.51 | 162  | 16    | 5.33  | 6.17 | 154 |
|    | [2]     | 14    | 12.38 | 2.88 |      | 14    | 6.18  | 5.15 |     |
| 2  | 30      | 23.36 | 8.18  |      | 30   | 18.34 | 9.23  |      |     |
|    | (1)     | 18    | 15.10 | 4.12 | 162  | 18    | 13.10 | 5.48 | 155 |
|    | (2)     | 12    | 8.25  | 5.19 |      | 12    | 5.24  | 4.66 |     |
| 3  | 20      | 10.63 | 5.49  |      | 20   | 8.50  | 6.07  |      |     |
|    | (1)     | 8     | 6.65  | 2.16 | 155  | 8     | 4.93  | 2.95 | 161 |
|    | (2)     | 6     | 2.32  | 2.59 |      | 6     | 2.37  | 2.62 |     |
|    | (3)     | 6     | 1.66  | 2.61 |      | 6     | 1.20  | 2.09 |     |
| 4  | 20      | 13.83 | 5.17  |      | 20   | 10.90 | 6.49  |      |     |
|    | (1)     | 7     | 6.64  | 1.30 | 135  | 7     | 5.74  | 2.39 | 145 |
|    | (2)     | 13    | 7.19  | 4.63 |      | 13    | 5.16  | 5.10 |     |
| 5  | 20      | 6.95  | 4.96  |      | 20   | 2.59  | 3.64  |      |     |
|    | (1)     | 4     | 3.24  | 1.34 | 21   | 4     | 1.19  | 1.86 | 27  |
|    | (2)     | 6     | 2.48  | 2.34 |      | 8     | 1.11  | 1.45 |     |
|    | (3)     | 10    | 1.24  | 2.21 |      | 8     | 0.30  | 1.35 |     |

(N と表示) とともに、表2に示した。

選択問題で指定以上の問題に解答した受験者の結果は、すべて含めて計算されている。この表から、数Ⅱ・Bの第3問(2)を除き、マークシートの得点の平均値が記述形式のそれを上回っていることがわかる(数Ⅰ・A, 第1問 [1] (2) も見かけ上そうになっているが、この部分は満点が異なるので比較はできない)。

また、マークシートの得点と記述の得点は、ほぼ対応して上下している。すなわち、マークシートで難しい問題は、記述でも難

しいと言える。

一方、標準偏差を見ると、おおむね記述形式の方がマークシート形式より大きい。このことは、名古屋大学レベルの大学において、受験者間の能力差の識別には、(問題内容は同等であっても) マークシート形式より記述形式による方が適切であることを示唆している。

5.2 マークシートと記述の相関関係

各受験者の2つの科目におけるマークシートの得点と記述の得点を合計し、それぞれ、マークシート得点と記述得点とした。

マークシート得点を横軸に、記述得点を横軸にとった散布図が図2である。

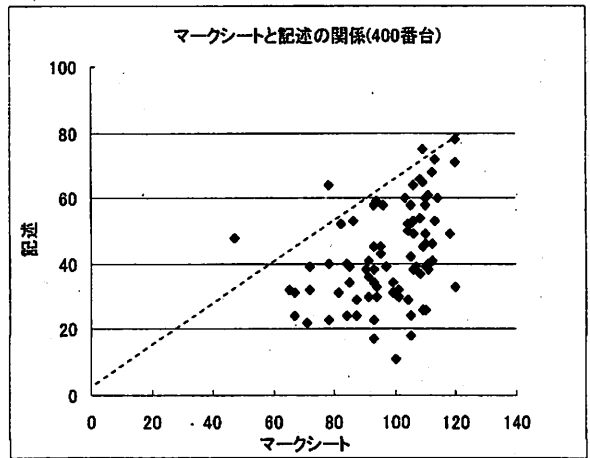
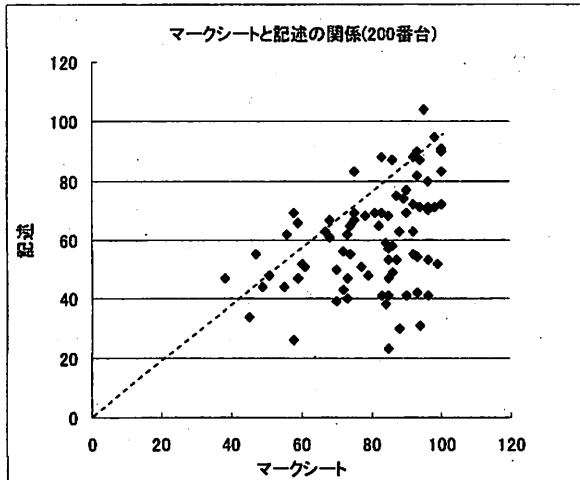
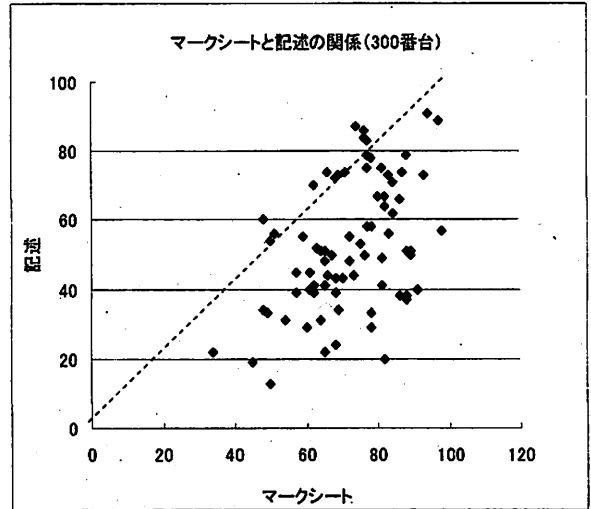
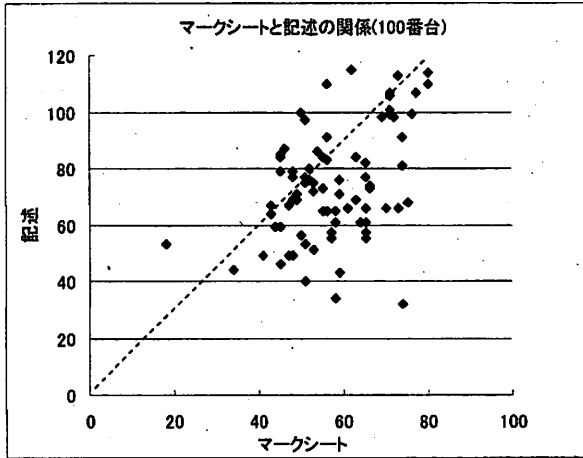


図2 マークシート形式と記述形式の間の散布図 (問題の組み合わせごと)

200番台と300番台の受験者についてはマークシート、記述とも100点満点となるが、100番台では、マークシートが80点満点、記述が120点満点であり、400番台では、マークシートが120点満点、記述が80点満点である。なお、満点を超える受験者があるのは、選択問題の両方に解答して得点している者に、現時点ではともに得点を与えて総点を求めているからである。

散布図の様相は、群によっていくらか違いが見られるが、おおよそ森田(2003)の見出したように、0点と満点を結ぶ線の右下に点が集中する結果となっている。ただし、両方の得点が0に近いという左下隅にはほとんど点がない。これは、受験者がすべて名古屋大学の合格者(入学者)であって、著しく低学力の者は存在しないこと、本人の処遇にはかかわりのないモニター調査であるにもかかわらず、全員が相当のモチベーションをもって取り組んだことを示すものであろう。

この結果は、前述のようにマークシート形式が全体として記述形式よりもやさしいという事実と整合的であるが、2つの試験の内容の違いを表すものであるかどうかは、従来の結果と同様、これだけから即断することはできないであろう。

本稿の段階では、試験結果の分析はきわめて初歩的な段階にとどまっている。今後、記述式答案の分析、受験者の学習観、学習方略との関係の分析、受験者の実際の入試成績と入学後の成績との関係の分析に取り組む必要がある。特に、受験者の学習観、学習方略との関係の分析について言えば、相対的にマークシート形式において得点の高い受験者が「暗記・再生」型の学習観をもち、記述形式において得点の高い受験者が「理解・思考」型の学習観をもつ傾向があることは見出されているが、このことにより明確な証拠を見出す必要がある。

#### 文 献

- 森田康夫(2003) センター試験と二次試験の「数学」の得点の相関について 大学入試研究ジャーナル No. 13, 95-97.  
西田吾郎(2003) 森田論文へのコメント 大学入試研究ジャーナル No. 13, 98.  
桜井捷海・藤井光昭・岩坪秀一・伊藤 圭・松田稔樹(2003) 筆記回答方式とマー

クシート方式により測定された学力の比較 大学入試研究ジャーナル No. 13, 63-69.

付記：本研究は文部科学省「先導的・大学改革推進委託事業」にもとづく「受験生の思考力、表現力等の判定やアドミッションポリシーを踏まえた入試の個性化に関する調査研究」の一部として実施された。