

試験科目としての小論文と総合問題の妥当性と信頼性

天野 哲彦, 樽松 理樹 (岩手県立大学)

課題発見力, 思考力, 表現力などを測るために, 教科テストとは異なる, 小論文, 総合問題, 面接といった評価方法を導入する大学は多い。そのことは暗記に頼らない多面的・総合的評価という観点から肯定すべきである。しかし, 主観評価が含まれることから, テストとしての信頼性確保に限界がないか, 公正な選抜が実施できているのか, 基礎学力は担保できているのか, といった疑念は残る。本稿では, 一見トレードオフにも見える両観点を両立させることはできるのかという課題について, 本学 A 学部での学校推薦型選抜改革として実施した小論文から総合問題への変更に基づく検証資料を提供する。

小論文から総合問題への変更は, 基礎学力の担保 (テストの妥当性確保) とテストの信頼性確保という観点から一定の成果が得られたと考える。また, 図表の読み取りを課す出題は, 言語能力と数理能力とにまたがる幅広い学力を測ることのできる可能性が示唆された。

キーワード: 学校推薦型選抜, 小論文, 総合問題, 妥当性, 信頼性, 識別力, 作題ガイドライン

1 はじめに

本学 A 学部は, 学校推薦型選抜において, 小論文 100 点満点と面接 50 点満点という, 小論文のウェートの大きな選抜を実施してきた。

この選抜方法に関しては, 読解力, 思考力, 表現力といった能力を総合的に判定できるという評価がある一方で, 基礎学力を担保できないのではないかと, 判定の信頼性に問題があるのではないかと, といった疑念も提起されてきた。

基礎学力の担保という点に関しては, 出願資格として全体の学習成績の状況 (旧評定平均値) 4.0 以上を課してはいるものの, 出願資格にすぎないために受験者間の学力比較には利用できない。仮に点数化したとしても微差を拡大して評価せざるを得ず, 識別力は低下する。また, 学校間の学力格差を考慮することもできない。したがって, 全体の学習成績の状況 (旧評定平均値) 4.0 という数値から, 「高校のカリキュラムを平均以上で修了した」という証明以上の意味を受け取ることは難しい。

判定の信頼性への疑念については, 複数の出願者を出している高校から, 高校での評価とは逆の合否判定になっているという意見が寄せられることもあった。

以上の背景から, 基礎学力の担保とテストの信頼性という 2 つの課題への対応として, 本学 A 学部の学校推薦型選抜を, 令和 5 年度選抜より小論文から総合問題へと変更して実施した。本稿では, この変更が上記課題の解決に寄与しているのかを, 令和 2 年度選抜から令和 4 年度選抜で実施された小論文と, 令和 5 年度選抜で実施した総合問題の結果を比較することで考

察する。本内容は, 同様の課題を抱えている大学に有用な情報になると考える。

2 小論文の分析

A 学部の小論文は, A4 判 2 ページ弱の課題文をもとにして, 初めに漢字の書き取り, 次に 100 字から 200 字での読み取り説明問題を課し, 最後に 600 字程度の自由論述を求めるという出題構成だった。最後の自由論述問題が配点の 7 割と大きなウェートを占めている。また, 自由論述問題以外は比較的平易な設問である。

前述の「高校での評価とは逆の判定になっている」との意見について, 小論文の得点の散らばりが小さいために, 小さな差違を重視せざるを得ず, そのことが合否の食い違いを生み, 受験生を推薦する高校側に納得を得られにくかった, というのが筆者の予想であった。つまり, 高校教諭は受け持つ高校生を総合的・多面的, かつ長期にわたって身近で見て, その実力を評価するのに対して, 採点者側は小論文という 1 科目, とくに自由論述で判定せざるを得ない。よって, このような選抜方法では, 採点者は答案の良否の差をつけることが難しいために, 慎重に評価することになり, 合格ライン付近には多くの受験生が密集して, 十分な識別ができていないのではないかと予想した。また, 1 科目での合否判定となると, そのときの出題テーマや出題形式によって得点に大きな揺れが生じて, テストとしての信頼性に欠けてしまうという問題もある。

以上の点について検証するために, A 学部学校推薦型選抜における令和 2 年度から令和 4 年度までの小

論文の得点について分析を加えた。表1に3年分の統計量を、図1に3年分の得点分布を示す。

表1 小論文の得点の統計量

	令和2年度	令和3年度	令和4年度
受験者数	60	64	61
最高点	83	89	99
最低点	21	46	20
最高-最低	62	43	79
中央値	53.0	67.0	55.0
平均値	50.7	67.8	56.6
標準偏差	14.9	10.2	16.5

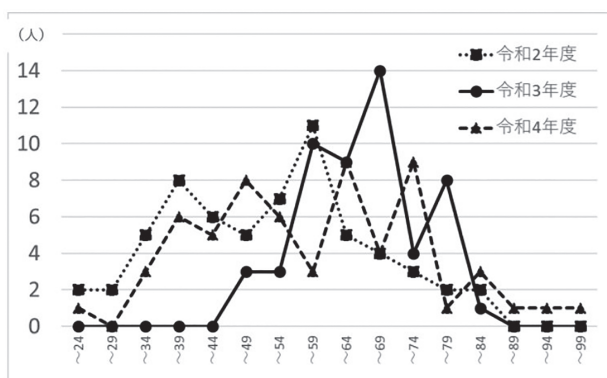


図1 小論文の得点分布

上記の3回の小論文試験の得点分布では、得点にある程度の散らばりが確認できる。このことから、得点差が小さいために、適切な選抜が実現しにくいのではないかという事前予想は外れた。

しかし、年度によって最高点と最低点の幅や平均値、中央値に大きな違いがあることから、出題された課題文、設問、採点者によって得点に変化が生じていたことが推測できる。また、令和4年度入試では、満点に近い高得点者が数人出ているが、自由論述問題で満点かそれに近い採点結果が多く見られるということは、一般的には考えにくい。評価には十分な差が出ているが、その評価が適切なものか、という別の課題があるとの見方もできる。

このような採点の信頼性という課題については、詳細なルーブリックや採点基準を作り、複数採点者間で採点調整をすることが取り得る対応策である。しかし、実際の運用は難しい。

例えば、ルーブリックや採点基準については、次のような項目を挙げることができる。

- (1) 表記・表現など表現形式
- (2) 文章構成（序論・本論・結論など）
- (3) 課題文や設問との応答
- (4) 独自の論述テーマ設定
- (5) 適切な具体性
- (6) 論拠や論理展開の説得力

ある程度小論文の練習をしている受験生であれば、(1)(2)(3)はほぼ達成でき、(4)は形としてできていることがほとんどである。そのため、採点者は(4)の「独自」、(5)の「具体性」、(6)の「説得力」といった主観的で曖昧な評価を重視せざるを得ない。逆に言うと、そのような採点基準を作っておかないと、得点差のある採点結果が出ない。このことは、得点分布を広げてメリハリのある採点結果をもたらすかもしれないが、ルーブリックや採点基準を作り、できるだけ客観的かつ公正な採点を実現しようとした意義自体が薄れてしまう。

また、A学部の場合も複数採点者間で話し合いながら採点調整を行っているが、短い採点時間の中での調整には限界がある。複数採点者の採点結果の平均を取るという方法は問題を緩和はしても、解決するとは言えない。採点者を3人、4人と増やすことも人的資源の限界から困難である。

3 総合問題の分析

3.1 総合問題の出題意図

上述のような、選抜方法としての小論文が抱える、信頼性、公平性、客観性の問題に対応するためと、基礎学力の適切な判定（妥当性の確保）という目的のために、A学部では、令和5年度学校推薦型選抜から、小論文に替えて総合問題を導入した。

これは、基礎学力を客観的に判定することを目的とするテストであるため、『令和5年度大学入学者選抜実施要項について（通知）』（文部科学省、2022）において「各教科・科目に係るテスト」は2月1日以降に実施すべきとされていることとの関係が問題になる。

対応としては、第一に、出題範囲を高校2年生の前半までに履修し、普通高校のみではなく、専門・総合高校でも等しく履修する範囲とした。なぜなら、この制度の趣旨は、受験競争の早期化・激化によって高校のカリキュラムを乱すことを防ぐことと、高校毎に履修教科が異なることからくる不公平を避けることにあると考えるからである。

第二に、特定教科の知識を測るよりも、読解力、分析力、表現力を問う問題作成を目指した。そのため、比較的分量のある文章や図表を資料として使い、記述

を求める設問を設けた。

総合問題は、1. 国語（現代文）、2. 図表読み取り、3. 数学（数学Ⅰ・数学Aまでの範囲）という3つの大問で構成し、それぞれ50点満点で計150点満点とした。大問1、2、3はそれぞれ表2に示す能力を測る5つから7つの問によって構成されている。また学校推薦型選抜全体の配点としては、総合問題と面接50点満点を合わせた200点満点とした。なお、調査書、志望理由書、推薦書などの事前提出書類は、面接の際の参考資料とした。

表2 総合問題の構成

大問1（国語、現代文の内容）	
言語的な知識	漢字の読み・書き取り、語句・ことわざの意味説明など
言語的な情報把握力	細かい情報も含め、文章を正確に読み取る
言語的な内容理解力	文章全体を理解し、解釈する
言語的な推論（仮説的推論）力	文章の理解・解釈を前提として、推測、評価、判断を行う
大問2（図表読み取り）	
数量的な情報把握力	数値も含め、資料を正確に読み取る
数量的な内容理解（帰納的推論）力	資料から特徴や傾向を読み取る
言語的な推論（仮説的推論）力	図表の読み取りを前提として、推測、評価、判断を行う
大問3（数学、数学Ⅰ・数学Aまでの範囲）	
数学的な知識・技能	定理・公式の提示、単純な計算など
数学的な情報処理力	資料や設問に従って、数式や数値を正しく処理する
数学的な推論（演繹的推論）力	数式や数値を操って推論する。証明問題など

表2で示した能力を考えるにあたっては、大学入試センター研究開発部による『「言語運用能力・数理分析力」試験の枠組みおよび作題ガイドライン』（大学入試センター研究開発部、2019）を参考にした。この「作題ガイドライン」では、測る能力を次のように分類している。

言語運用力

- L1：情報の把握
- L2：内容の理解
- L3：推論と推察

数理分析力

- M1：数理的な表現・原理の理解

M2：ルール・法則性の理解と適用

M3：資料からの情報抽出・分析

M4：帰納的・演繹的推論の適用

A学部での出題は、大学入試センター研究開発部による作題例と比べると、英語による出題がなく、現代文と数学（数学Ⅰ・数学A）という高校教科の内容に沿ったものとなっている。また、大問2は主に図や表といった資料の読解と解釈を求める出題であり、大学入試センター研究開発部「作題ガイドライン」にある「数理分析力」の枠を越えた領域まで扱う。そのため、上記のように、1. 国語（現代文）、2. 図表読み取り、3. 数学（数学Ⅰ・数学Aまでの範囲）という3つの領域でそれぞれ測る能力を定義することとした。また、推論能力については、「帰納的推論」「演繹的推論」「仮説的推論」の3種に分類した。そして、「帰納的推論」（特徴や傾向の読み取り）は主に大問2（図表読み取り）で測ることができる能力、「演繹的推論」は大問3（数学）で正確に測ることができる能力として位置づけた。また、「仮説的推論」は大問1（国語）と大問2（図表読み取り）で測ることができ、主観を交えた推論が混じることを前提としている。

3つの領域で測る能力はそれぞれの表2の下に行くほどより高度で総合的な思考力・表現力を必要とするものとなっており、設問の配列も基本的にはその順番となっている。

全体として、大問3（数学）も含めて、資料や文章を読み取ることを前提とする出題であり、読解力を重視している。また、大問3（数学）では、数式による証明や計算だけでなく、言葉や図示による説明も可能な問題も出題した。

また、3つの大問すべてに記述式問題が含まれている。しかし、自由な意見論述の出題はなく、意図的に記述の方向を絞っている。これは、基礎学力確認というテストの目的と、主観評価の縮減という出題意図とによる。

3.2 総合問題の採点結果

令和5年度入試の総合問題の受験者数は61人であった。採点結果について、表3に得点の統計量、図2に合計点の得点分布、図3から図5に各大問の得点分布を示す。

表3 総合問題の得点の統計量

	合計	大問1	大問2	大問3
最高点	118	43	45	43
最低点	32	15	6	3
中央値	83.0	34.0	31.0	18.0
平均値	81.3	32.1	30.1	19.1
標準偏差	15.3	5.9	8.4	8.7

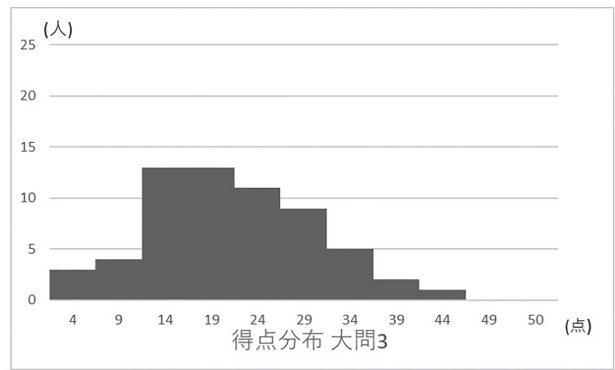


図5 総合問題の大問3（数学）の得点分布

大問3つの合計点は、ほぼ左右対称の分布となっている。大問単位で見ると、大問1（国語）は、分布が高得点に偏り、散らばりが小さいのに対して、大問3（数学）は、分布が低得点に偏り、散らばりが大きい。このような大問間の差は、来年度以降の出題で改善の努力が必要である。

次に各大問間の関係について分析する。

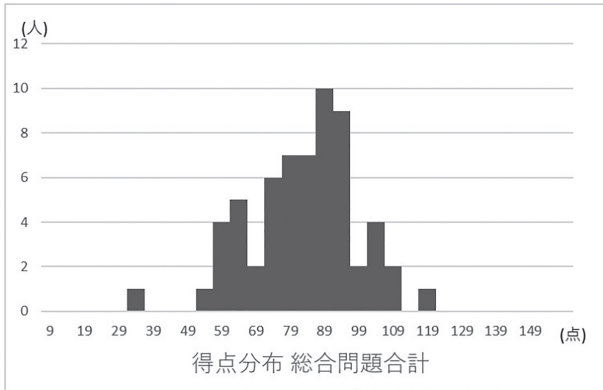


図2 総合問題の合計点の得点分布

3.3 大問間の関係

図6から図8に大問間の散布図及び相関係数を示す。また図9に合計と各大問間の共分散比を示す。

共分散比は次の式で表され、いくつかの変数の合計得点を求めた際の、合計得点に対する各変数の寄与の大きさを表す指標である。

$$\text{共分散比} = \frac{\text{ある科目の得点と総合得点との相関係数} \times \text{ある科目の標準偏差}}{\text{合計得点の標準偏差}}$$

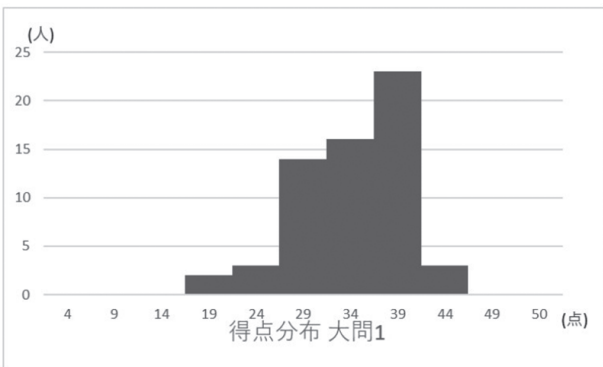


図3 総合問題の大問1（国語）の得点分布

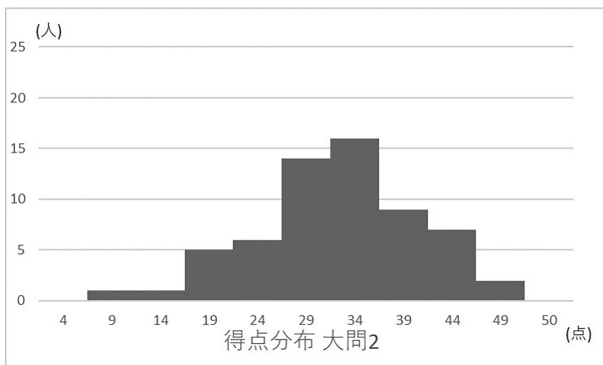


図4 総合問題の大問2（図表読み取り）の得点分布

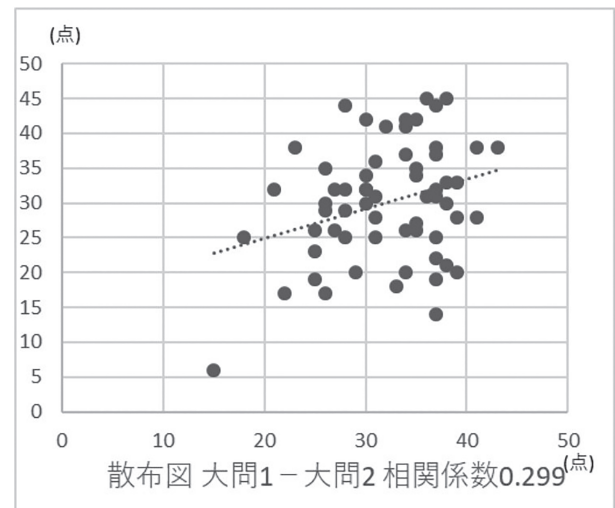


図6 大問1と大問2の散布図

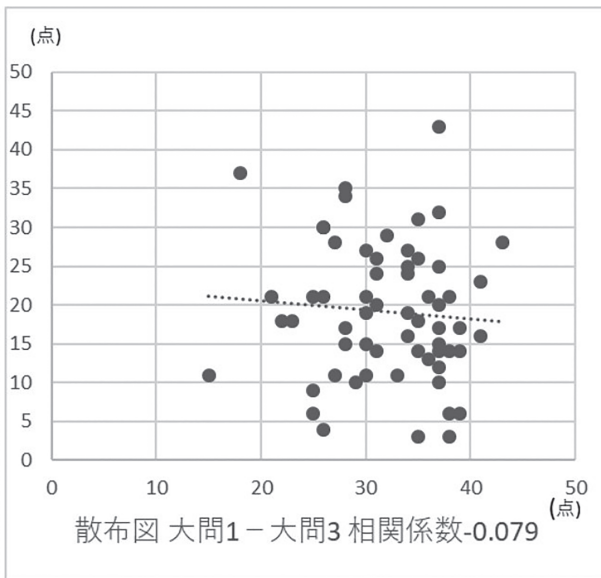


図7 大問1と大問3の散布図

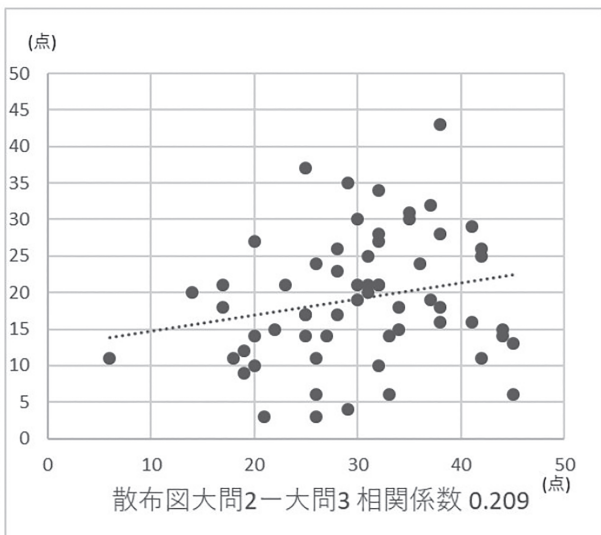


図8 大問2と大問3の散布図

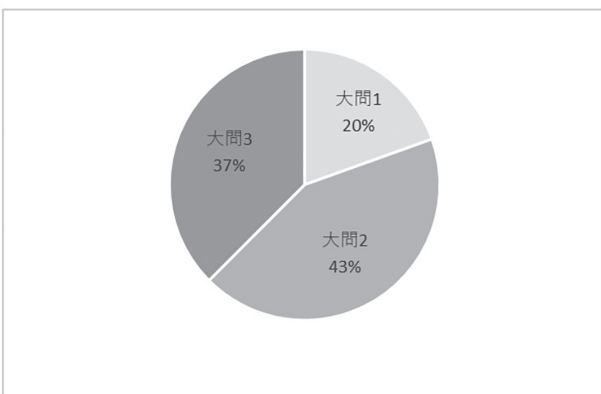


図9 総合問題合計と各問間との共分散比

図6から図8より、大問1(国語)と大問2(図表読み取り)、大問2(図表読み取り)と大問3(数学)との間には正の相関があるのに対し、大問1(国語)と大問3(数学)の間には弱い負の相関がある。1回だけの試験で確定的なことは言えないが、本結果から、言語的能力と数理的能力の間に相関は無いと考えられる。今回の出題では大問3(数学)は単なる計算や数学的な証明ではなく、比較的長い資料文による説明がなされているにも関わらず、このような結果が得られている。このことから、小論文という言語的能力を測るテストのみでは、基礎学力を適切に測ることはできない可能性があることがわかる。言語的能力と数理的能力という2本の柱を立てることは基礎学力の測定のために必要と言える。

他方、大問2(図表読み取り)は、大問1(国語)と大問3(数学)の両方と相関があることから、多様な分野からの出題が困難な場合は、図表読み取りを中心とした出題とすることによって、比較的バランス良く基礎学力を測ることができると考えられる。このことは、図9より大問2の共分散比が大きいことからからも言えよう。

3.4 主観の混じる設問(仮説的推論)について

令和5年度入試の総合問題において、与えられた文章や図表に含まれる情報以外の知識や考えを導入して解答することを求める設問は、大問2(図表読み取り)の最後の設問である問5のみであり、設問の内容は「男女共同参画の現状について、すべての図表から読み取れることを250字から300字で答えなさい」(15点満点)である。この問では、意見論述が求められているわけではないが、5つある図表に表れている傾向を一貫したストーリーのもとに説明するには、帰納的推論に加えて、自分なりのものの見方に基づいて理由を推論する、仮説的推論が必要となる。その意味で解答者の主観が混じることが避けられない設問である。

採点結果は図10に示すとおりであり、得点分布には大きな散らばりが出ている。しかし、図11に示す箱ひげ図(総合問題合計点の低い順から取った5グループを横軸に、各グループの大問2問5の得点分布を縦軸に取った)を見ると、平均値(×印)と中央値(一印)はゆるい右肩上がりであり、総得点と大問2問5の得点との相関係数は0.29と小さな正の相関となっている。

このことは、この設問のような自分が予め持っている知識や考え方を使い、主観も背景に持ちながら記述することを求める出題が、他の主観を交えずに知識や

論理で解答できる出題とは異なる能力を測定している
と見ることができ、また採点者の採点が解答者の学力
を正確に評価できていないとも見ることが出来る。

また、この設問は自由論述問題ではないにも関わら
ず、例えば政策提言など自分の意見を述べている答案
も多く見られた。このことから小論文対策を積んだ受
験生が、設問の意図を誤解してしまったことが、得点
分布と総合的な学力との関係を不明瞭にしていると思
えることも出来る。

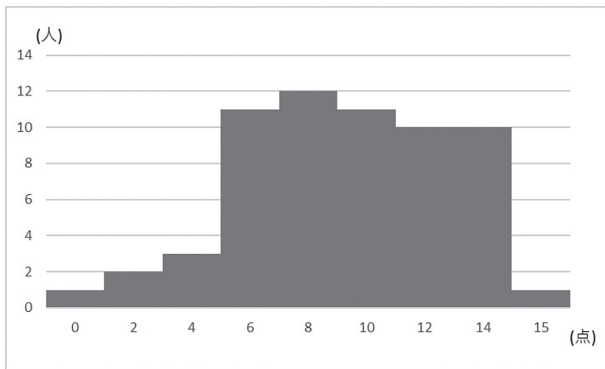


図10 総合問題大問2問5の得点分布

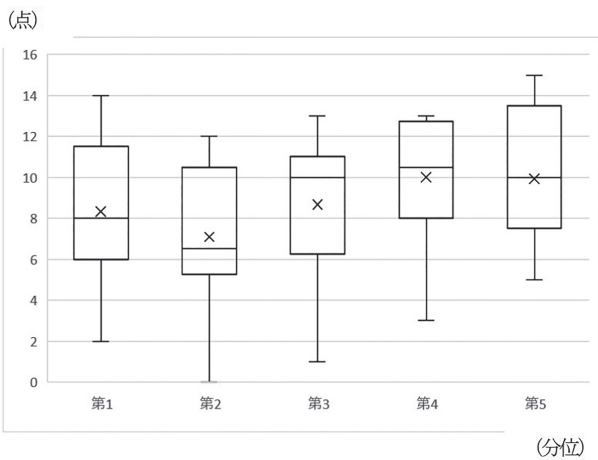


図11 総合問題計－大問2問5得点箱ひげ図

以上のように、総合問題の得点には複数の要因が関
わっている可能性もあり、他の事例も取り入れながら、
分析を進める必要がある。しかし、主観を交えて論述
するという方向に進むことが、基礎学力を適正に評価
するという目的から外れてしまう危険があることをこ
の事例は示唆している。

4 終わりに

小論文から総合問題へと選抜方法を変更したことによ
って、基礎学力（国語と数学の学力の他、図表や文

章をもとに推論する力）の担保（テストの妥当性確保）
と、主観評価を縮減することで採点の信頼性を高め
るという目的は、一定程度達することができたと言え
る。基礎学力の確保については、入学者の成績追跡や
次回入試の結果を踏まえてさらなる分析が必要であ
る。

本学 A 学部で令和 5 年度に導入した総合問題にお
ける、国語、図表読み取り、数学という 3 分野の成績
の関係から、基礎学力確保のためには言語能力分野と
数理能力分野の両分野の出題が必要なことと、図表読
み取りは両分野にまたがる能力を測ることが出来る可
能性が示唆された。また、図表などの資料から正確に
傾向を読み取り、仮説を推論することは、データサイ
エンスで統計学や情報技術を駆使する前の重要な能力
となることから、教育や入試の場で今後さらに重視す
べきである。

テストの信頼性確保という観点からは、主観を交え
た記述や採点の領域を縮減するという目的は総合問題
への変更によって達成できた。しかし、自由な発想や
表現は今後さらに社会的な重要度が増す。形の整った
文章を書くという点ではすでに AI が人の能力を凌駕
しつつあるので、入試においては、ユニークさや独自
の発想・表現を評価する領域を確保する必要がある。
そのためには、主観評価が必要なテストの採点基準や
ルーブリックのブラッシュアップに加え、採点者の訓
練、採点システムの改善を考えるべきである。形式は
AI が評価し、形式知となり得ない観点から人が評価
するという分業が必要となる。

参考文献

文部科学省 (2022), 「令和 5 年度大学入学者選抜実施要項 (通
知)」, 4 文科高第 302 号, 令和 4 年 6 月 3 日, [https://
www.mext.go.jp/content/20210617-mxt_
daigakuc02-000010813_1.pdf](https://www.mext.go.jp/content/20210617-mxt_daigakuc02-000010813_1.pdf) (2023 年 4 月 23 日アクセス)
大学入試センター研究開発部 (編) (2019), 「『言語運用力・
数理分析力』試験の枠組みおよび作題ガイドライン
(Ver.2)」『大学での学修に必要な基本的学力の評価方法に
関する研究』報告書, 127 - 172