

高校教員は新傾向の国語記述式問題をどう捉えているか？

—全国調査から—

宮本 友弘, 倉元 直樹, 庄司 強 (東北大学)

全国の高校 2,000 校の教員を対象に新共通テストの国語で導入予定である新傾向の記述式問題（イメージ例, モデル問題例）に対する質問紙調査を実施した。その結果, 新傾向の記述式問題は, マーク式及び旧来の記述式問題よりも, 新奇性が高く評価された。また, 解答に必要なとされる資質・能力の判定においては, 国語の新傾向の記述式問題は, 国語のマーク式及び旧来の記述式問題よりも, 数学の試験問題に類似していた。こうした新傾向の記述式問題が持つ特異性の認知においては, 教員と高校生は同様の傾向にあることが示唆された。

1 問題

2020 年度から実施される「大学入学共通テスト」（以下, 新共通テスト）では, 国語において記述式問題が出題される。そのねらいは「複数の情報を統合し構造化して新しい考えをまとめる能力やその過程や結果を表現する能力」（高大接続システム改革会議, 2016）を適切に評価することにある。背景には, 大学入試センター試験はマークシート式であること, また, 個別学力試験はたとえ記述式を実施していたとしても, こうした能力を十分に評価していないとする批判がある。

このため, 上記能力を適切に評価できる記述式問題の整備が急務となり, そのプロトタイプとして, 「大学入学希望者学力評価テスト（仮称）」のイメージ例（2015 年）, 「大学入学共通テスト（仮称）」記述式問題のモデル問題例（2017 年）が公表されてきた。それらにおいては, 題材として従来の評論や小説だけでなく, 日常生活でみかける文章, 会話文, 図表等が使用され, また, 他者の考えを推論させたり, 比較させたりするような設問がなされるなど, これまでにない意匠が凝らされている。

次のステップとしては, 当然ながら, こうした新傾向の記述式問題のテストとしての評価が求められる。その際, 利用者である高校生や教員らの反応を実証的に検証することは不可欠である。なぜなら, テストの開発者が規定する測定内容は, それだけでは表面的妥当性（face validity）を示すにすぎないからである（田中・宮本・倉元, 2018）。また, 新傾向の記述式問題のそもそもの開発動機が, 上記した通り, 既存のセンター試験や個別学力試験の「不具合」にあるというならば, それらとの異同も明らかにすべきであろう。

以上の問題意識のもと, 筆者らは, 平成 28 年度には高校生 1,129 名を対象にイメージ例を使用して, ま

た, 平成 29 年度には高校生 1,535 名を対象にモデル問題例を使用して, 大学入試センター試験, 個別学力試験（記述式問題）との比較調査を行ってきた（表 1 参照）。その際, 試験問題に解答してもらうとともに, 各試験問題に対する印象評価や解答に必要な資質・能力の判定といった意識調査も行った。また, 比較対象として数学でも実施した。その結果, 印象評価においては, イメージ例は, 大学入試センター試験や個別学力試験よりも「新奇性」が高く評価された（宮本・倉元・田中, 2017）。また, 解答に必要な資質・能力の判定では, イメージ例は, 大学入試センター試験, 個別学力試験とは異なり, むしろ, 数学の試験問題に近かった（田中・宮本・倉元, 2018）。

このように新傾向の記述式問題を高校生がどのように捉えているかの一端が明らかになりつつある。本研究では, さらに, 利用者側の視点に立った新傾向の記述式問題に対する評価を進めるために, 対象を高校教員に広げ同様の調査を行った。

2 方法

2.1 調査対象

「平成 29 年度版 全国高等学校一覧」（全国高等学校長協会, 2017）を抽出台帳とし, 無作為に 2,000 校を抽出した。

2.2 実施期間

調査は 2018 年 2 月～3 月にかけて実施された。

2.3 調査内容

2.3.1 試験問題

評価対象となる試験問題は, 先述した高校生対象の調査で使用した試験問題をそのまま使用した。年度と

教科から計4冊子あり、各冊子の試験問題は、①マーク式問題(大学入試センター試験)、②旧来の記述式問題(個別学力試験)、③新傾向の記述式問題の3題から構成された。各試験問題の出典と題材等は、表1の通りであった。

2.3.2 質問紙

試験問題を参照しながら回答する質問紙は、冊子の内容に応じて4種類作成した。質問内容は共通であり、次の通りであった。

①各問題に対する印象評価 これまでの高校生対象の調査で使用したSD法(Semantic Differential method)による評定尺度12項目を、教員向けに一部修正して使用した(表4参照)。例えば、「無意味な—有意義な」では、教員という立場からの視点を明確にするために、「生徒にとって」を付記した。問題ごとに、-2~+2の5段階で評定してもらった。

②各問題の解答に必要な資質・能力の判定 高校生対象の調査で使用した「新しい時代にふさわしい高大接続の実現に向けた高等学校教育、大学教育、大学入学選抜の一体的改革について～すべての若者が夢や目標を芽吹かせ、未来に花開かせるために～(答申)」(中央教育審議会, 2014)と「高大接続改革の進捗状況について」(文部科学省, 2016)を参考に作成した22項目をそのまま使用した(表5参照)。問題ごとに各資質・能力が必要かどうかを判定してもらった。

③回答者の属性や個別学力試験に対する意識に関する質問 勤務校の設置者と大学入試センター試験の受験状況、教員歴、校務分掌での進路指導部担当歴、個別学力試験に対する意識等について尋ねた。

2.4 手続き

調査は郵送法で実施された。年度と同じ国語と数学

の冊子及びそれらに対応した質問紙をワンセットとし、2セット作成した。調査対象校には、どちらかの年度のセットを無作為に割り当て送付し、各教科1名の教員に回答するよう依頼した。その際、回答は任意とする旨の説明をアンケートの表紙に明記した。回答終了後は、同封した返送用封筒でアンケート用紙を返送してもらった。なお、本調査は、東北大学高度教養教育・学生支援機構研究倫理委員会の承認(平成29年12月25日)を受けて実施した。

3 結果

3.1 サンプルの特徴

3.1.1 回収状況

H28版の国語は561名(回収率56.1%)、数学は577名(回収率57.7%)、H29版の国語は573名(回収率57.3%)、数学は598名(回収率59.8%)、総計2,309名(回収率57.5%)から回答が得られた。なお、同じ年度でも教科で人数が異なるのは、学校によってはどちらか一方の教科しか返送されなかったためであった。

3.1.2 属性

①勤務校の状況 設置者については、全体では、国立1.2%、公立71.5%、私立26.9%、その他・無回答0.4%であった。この比率は、年度・教科別でも同様であった。なお、大学入試センター試験の受験状況の回答には多くの不備が見られたので分析から除外した。

②回答者の教員歴 平均18.4年($SD=10.8$)であり、年度・教科別に見ても著しい差は認められなかった。

③進路指導部担当歴 全体では78.9%が進路指導部を経験し、のべ年数は平均6.9年($SD=5.9$)であった。年度・教科別に見ても同様であった。

以上から、年度・教科によって教員としての基本的な属性の分布に著しい違いはないとみなせた。

表1 各冊子の試験問題の出典と題材等

冊子		マーク (大学入試センター試験)	旧来 (個別学力試験)	新傾向
H28国語	出典	平成22年度本試験問題	平成23年度東北大学前期日程試験問題、	「大学入学希望者学力評価テスト(仮称)」で評価すべき能力と記述式問題イメージ例【たたき台】
	題材等	評論	小説	問題イメージ<例1>
H29国語	出典	平成22年度本試験問題	平成24年度東北大学前期日程試験問題	「大学入学共通テスト(仮称)」記述式問題のモデル問題例
	題材等	小説	評論	モデル問題例1
H28数学	出典	平成21年度追試験問題	平成14年度東北大学前期日程試験問題(文系)	「大学入学希望者学力評価テスト(仮称)」で評価すべき能力と記述式問題イメージ例【たたき台】
	題材等	三角関数	三角関数	問題イメージ<例4>
H29数学	出典	H28と同じ	H28と同じ	「大学入学共通テスト(仮称)」記述式問題のモデル問題例
	題材等	三角関数	三角関数	モデル問題例4

3.1.3 個別学力試験に対する意識

表2は、年度・学年別に、個別学力試験の検討への取り組みを集計した結果である。年度・学年によって有意な偏りは認められなかった。年度・教科にかかわらず「必要に応じて解いている」が約6割を占めた。

表3は、個別学力試験の改善の必要性を集計した結果である。有意な偏りが見られ($\chi^2(6)=63.57$, $p<.001$), 残差分析によれば、年度に関わらず、国語では「ある程度改善が必要である」が有意に多く、一方、数学では「現状のままで良い」が有意に多かった。

表2 個別学力試験の検討への取り組み (%)

	毎年自分で解いている	必要に応じて解いている	あまり分析していない
H28 国語	28.9	59.0	12.1
H28 数学	25.8	63.8	10.5
H29 国語	30.3	58.8	10.9
H29 数学	25.6	59.2	15.2

表3 個別学力試験の改善の必要性(%)

	現状のままでよい	ある程度改善が必要である	大いに改善が必要である
H28 国語	41.5	55.0	3.5
H28 数学	54.6	42.1	3.3
H29 国語	38.4	55.5	6.2
H29 数学	57.2	39.4	3.5

3.2 印象評価の分析

3.2.1 因子構造

試験問題をすべて込みにして2相データに置き換え、因子分析(主因子法)を行った。固有値の減衰状況から3因子が妥当と判断した。表4に示す通り、第1因子に負荷量の高い項目は、すべて「生徒」を付記して教員向けに修正した項目であった。教員の立場から生徒にとっての意義や有益性を評価する内容であることから「教育的価値」と命名した。第2因子は、教科の専門家としての目からみた問題のあり方を示唆する内容であることから「教科的品位」と命名した。第3因子は、問題の新しさや意匠の凝らし方を評価する内容であることから「新奇性」と命名した。

表4 因子分析の結果

(プロマックス回転後のパターン行列と因子間相関)

	I	II	III
12生徒にとって無意味な-生徒にとって有意義な	.896	.006	.093
10生徒の役に立たない-生徒の役に立つ	.859	.016	.127
9生徒に解かせたくない-生徒に解かせたい	.786	.089	.007
11生徒の実力がわからない-生徒の実力がわかる	.732	-.021	-.065
8生徒にはむりそうな-生徒にはできそうな	.515	-.125	-.234
1つまらない-おもしろい	-.053	.823	.334
2嫌いな-好きな	.007	.737	-.013
3ふざけた-まじめな	.015	.468	-.236
7下品な-上品な	.004	.452	-.161
5典型的な-奇抜な	.028	-.206	.826
4古い-新しい	.049	.189	.587
6素直な-意地悪な	-.169	-.286	.478
	因子間相関	I	-.567
		II	-.184
			-.023

3.2.2 印象評価の比較

各因子に負荷量の高い項目の評定値の平均を尺度得点とした。今回の調査で使用した計10個の試験問題

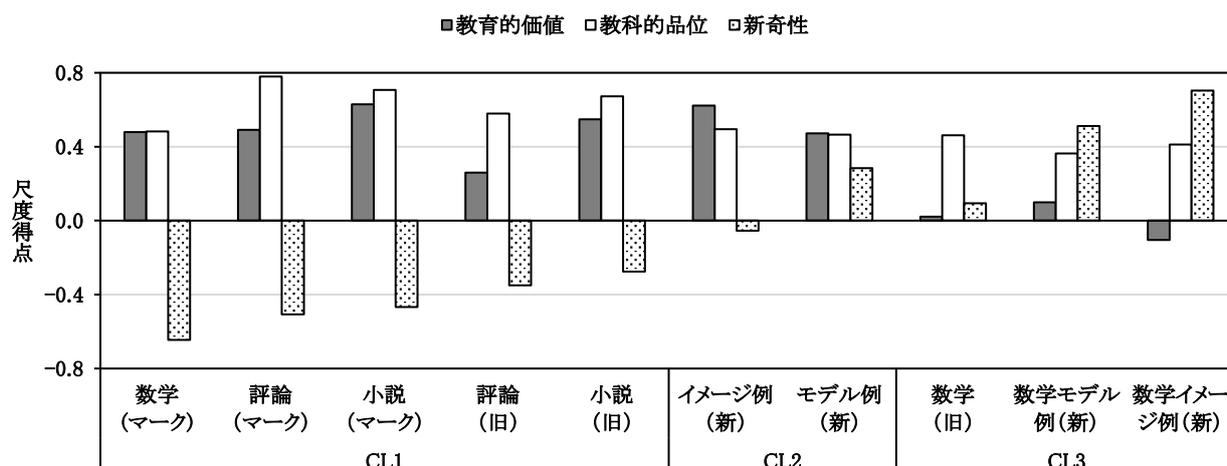


図1 クラスタ別各試験問題の尺度得点

表5 各資質・能力が試験問題の解答に必要と判定された割合(%)

項目	小説 (マーク)	評論 (マーク)	小説 (旧)	評論 (旧)	モデル例 (新)	イメージ例 (新)	数学 (マーク)	数学 (旧)	数学モデル 例(新)	数学イメ ージ例(新)
1言葉の働きや役割に関する理解(言葉の働き)	79.9	80.2	59.1	72.8	48.9	45.9	23.8	28.4	59.6	62.4
2言葉の特徴や決まりに関する理解(言葉の特徴)	68.6	62.9	45.7	57.8	38.6	38.2	25.5	29.6	45.1	51.6
3言葉の使い方に関する理解(言葉の使い方)	75.2	65.9	64.3	64.0	55.3	53.9	17.0	25.0	48.6	56.3
4言語文化に関する理解(言語文化)	36.3	32.9	37.0	45.2	22.9	11.1	3.1	3.8	26.5	23.4
5一般常識や社会規範(一般常識)	31.6	54.5	23.6	44.0	77.5	70.0	5.3	3.0	30.3	39.3
6情報を様々な見方でよく調べ、まとめる力(情報集約)	5.6	16.4	13.8	29.7	85.3	84.3	28.3	33.7	63.7	72.3
7自分の文の構成や表現形式を振り返り評価する力(文章評価力)	8.2	10.2	63.2	64.9	65.6	72.0	4.3	28.4	32.0	46.6
8言葉や数式によって感じたり想像したりする力(感受性)	85.0	24.6	85.4	28.1	29.5	23.2	35.3	50.8	55.9	62.9
9イメージを言葉や数式にする力(イメージ表現)	37.3	10.9	77.1	40.8	52.2	45.9	28.4	42.0	64.8	79.4
10言葉や数式を通じて伝える力(伝達力)	18.2	15.7	73.9	68.1	77.7	79.1	14.1	50.3	50.6	64.8
11考えを形成し深める力(思考形成)	23.9	52.1	52.3	73.6	63.2	64.3	31.3	54.9	51.1	64.8
12自分の感情をコントロールして学びに向かう態度(感情統制力)	31.4	24.1	27.3	27.7	26.2	18.8	10.4	11.6	16.4	19.1
13歴史の中で創造され継承されてきた言語文化の担い手としての自覚(歴史継承)	12.2	14.1	13.0	26.0	14.7	4.1	2.0	2.1	6.9	5.9
14粘り強く、柔軟に考え考察を深めていく態度(粘り強さ)	25.7	56.6	42.9	66.3	56.0	52.1	42.6	61.4	61.1	71.1
15多様な考えを生かし、問題解決する態度(多様な考え)	11.2	21.1	17.0	25.3	67.7	58.6	29.4	40.2	53.1	60.8
16読解力	86.0	96.3	81.6	92.0	67.2	56.3	35.0	34.9	84.3	86.5
17要約力	22.7	49.6	57.5	85.5	68.4	59.6	7.8	18.0	39.2	44.5
18表現力	19.9	10.5	81.2	71.2	75.6	75.9	8.3	39.7	45.1	65.0
19コミュニケーション力(コミュニケーション)	17.8	5.9	26.4	10.8	45.0	37.0	1.6	5.2	11.7	12.1
20統計的思考力	2.3	5.5	1.8	10.8	49.4	84.3	4.6	5.8	12.9	10.1
21論理的思考力	31.1	83.2	37.0	83.8	66.3	70.2	55.5	73.3	62.5	74.4
22図やグラフを描いたり読んだりする力(図やグラフ)	0.7	1.6	0.5	2.4	59.0	89.5	71.8	61.0	66.5	76.6

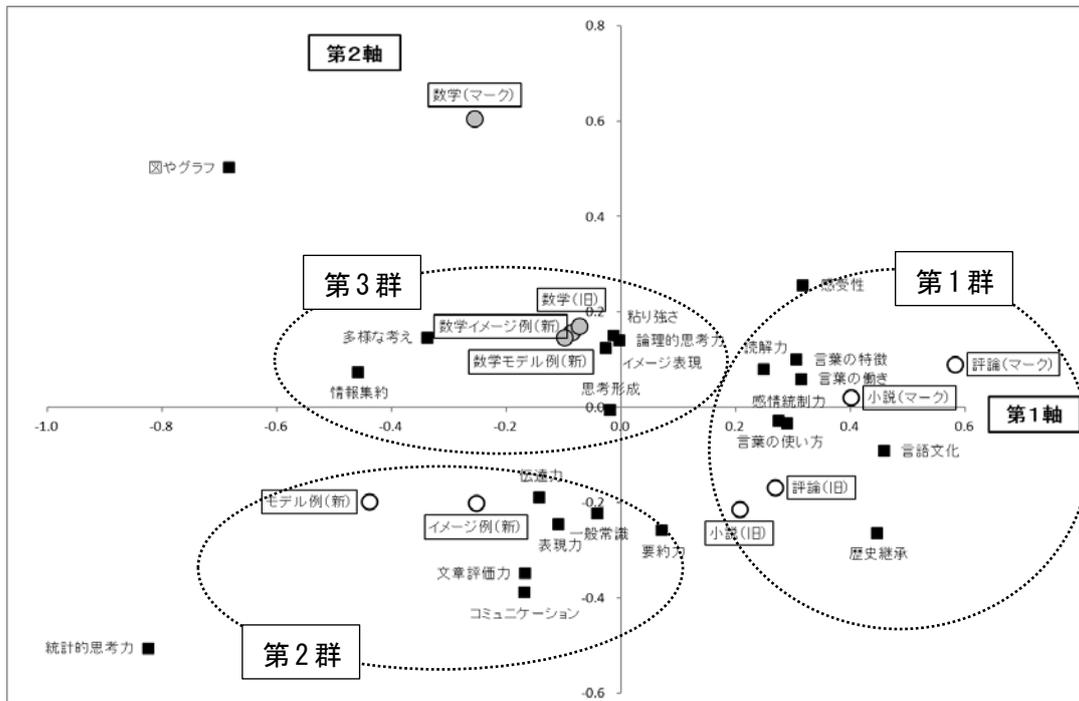


図2 資質・能力と各試験問題のマッピング

それぞれの3つの尺度得点(平均)を求め、それらを使ってクラスター分析(Ward法)を行った。デンドログラムから、3クラスター(以下、CL1, CL2, CL3)が妥当と判断された。図1は、クラスターごとに、各試験問題の尺度得点を示したものである。教科的品位はいずれのクラスターも同水準であった。CL1は、教育的価値がポジティブ、新奇性はネガティブであり、すべてのマークと、国語の旧来の記述式が分類された。CL2は、教育的価値はCL1と同等であるが、新奇性がCL1よりもポジティブであり、国語の新傾向の記述式からなった。CL3は教育的価値がCL1, CL2よりもネガティブである一方、新奇性がよりポジティブであり、数学の旧来の記述式及び新傾向が分類された。

3.3 解答に必要な「資質・能力」の分析

表5は、各資質・能力が、それぞれの試験問題の解答に必要と判定された割合を示したものである。50%以上(網掛け部分)に着目すると、国語、数学ともに、新傾向の記述式は、マークよりも、旧来の記述式と共通する資質・能力が多かった。また、両教科の新傾向どうしても共通する資質・能力が多かった。

さらに資質・能力と試験問題の関連を可視化して捉えるために、対応分析を行った。第1軸(イナーシャ=0.08814, 寄与率48.0%)を横軸、第2軸(イナーシャ=0.04496, 寄与率24.5%)を縦軸にして、資質・能力と試験問題を布置した。図2に示す通り、大きく次の3つの群に分かれた。第1群は、国語のマークと旧来の記述式と、言葉の知識や理解に関する資質・能力が布置された。第2群は、国語の新傾向の記述式と、言葉の運用に関する資質・能力が布置された。第3群は、マークを除く数学の問題と、深い学びに関する資質・能力が布置された。

3.4 個別学力試験に対する意識と新傾向の記述式の評価の関連

3.4.1 検討の取り組みとの関連

対象者を、個別学力試験を「毎年自分で解いている」(以下、H群)、「必要に応じて解いている」(以下、M群)、「あまり分析していない」(以下、L群)の3群に分け、新傾向の記述式に対する印象評価と、解答に必要な資質・能力の判定を比較した。印象については分散分析を、資質・能力の判定では χ^2 検定を行った。表6、表7には群間差が見られた結果を示した。下位検定の結果をみると、検討の取り組みの程度による新傾向の記述式に対する印象評価や、必要な資質・能力の判定に一貫した傾向は見出せなかった。

表6 検討の取り組みの群間差が見られた印象尺度

試験問題	尺度	多重比較の結果
国語 イメージ例	新奇性**	H(-0.19) < M(0.00), L(0.02)
数学 モデル例	教育的価値*	H(0.20), M(0.11) > L(-0.14)

注)カッコ内の数値は各群の尺度得点の平均

* $p < .05$ ** $p < .01$

表7 検討の取り組みの群間差が見られた資質・能力

試験問題	資質・能力	各群の選択者の割合(%)と 残差分析の結果		
		H	M	L
国語 イメージ例	5 一般常識*	75.9	69.4	58.8*▽
	9 イメージ表現*	37.7*▽	48.5	52.9
	18 表現力*	71.5	79.6*▲	67.7
数学 イメージ例	9 イメージ表現*	76.4	82.5*▲	70.0
	10 伝達力*	59.5	68.6*▲	55.0
	21 論理的思考力*	72.3	77.9*▲	61.7*▽

注)▲は期待度数よりも有意に高い、▽は期待度数よりも有意に低い

* $p < .05$

3.4.2 改善の必要性との関連

対象者を、個別学力試験の改善の必要性に対する回答のうち「現状のままでよい」を現状維持群、「ある程度改善が必要である」「大いに改善が必要である」を要改善群とする2群に分け、3.4.1と同様に、新傾向の記述式に対する印象評価と、解答に必要な資質・能力の判定を比較した。印象評価では t 検定を、資質・能力の判定ではFisherの直接法を行った。表8、表9には群間差が見られた結果を示した。

印象評価ではすべて要改善群が現状維持群よりも有意に高かった(表8)。また、有意差がみられた尺度としては、主に教育的価値、教科的品位であった。

資質・能力の判定では、国語モデル例の「言葉の特徴」(網掛け部分)以外は、すべて要改善群の選択率が有意に高かった(表9)。有意差が見られた資質・能力は、図2において、新傾向の記述式の近くに布置されているものが多い。

4 考察

高校教員の試験問題に対する印象評価は、高校生と同様に3因子であったが、意味内容には異同が見られた。「学ぶ立場」と「教える立場」では、試験問題の印象を捉える観点が異なるのは当然といえよう。唯一、負荷量の高い項目が1項目だけ異なったものの、「新奇性」という意味内容の因子は共通に見出された。この次元は、発達や経験に関わらず、頑健に存在するのかもしれない。また、試験問題間の印象評価を比較す

表 8 改善の必要性の群間差がみられた印象尺度

試験問題	尺度	現状維持	要改善	
国語	イメージ例	教育的価値*	0.54	0.69
		教科的品位*	0.42	0.55
	モデル例	教育的価値***	0.32	0.56
		教科的品位***	0.31	0.56
		新奇性***	0.19	0.35
数学	イメージ例	教科的品位***	0.31	0.53
		教育的価値**	0.00	0.22
	モデル例	教育的品位**	0.22	0.53

* $p < .05$ ** $p < .01$ *** $p < .001$

表 9 改善の必要性の群間差がみられた資質・能力

試験問題	尺度	現状維持	要改善	
国語	イメージ例	1 言葉の働き*	40.3	50.6
		5 一般常識**	62.8	75.3
		14 粘り強さ**	45.1	57.2
		17 要約力*	54.0	63.4
		18 表現力*	70.8	79.7
		22 図やグラフ*	85.4	91.9
		モデル例	2 言葉の特徴*	45.0
	9 イメージ表現*	46.0	56.9	
	10 伝達力*	73.0	80.8	
	14 粘り強さ*	49.3	60.8	
	15 多様な考え**	60.2	72.6	
数学	イメージ例	6 情報集約**	67.8	78.0
		8 感受性**	57.6	69.1
		9 イメージ表現***	73.6	86.9
		10 伝達力**	58.8	71.8
		14 粘り強さ*	66.6	76.4
		15 多様な考え方*	56.9	66.0
		17 要約力**	38.6	52.9
		18 表現力***	58.5	73.4
		21 論理的思考力**	70.1	80.7
		モデル例	9 イメージ表現*	60.7
	10 伝達力*	46.5	56.9	
	15 多様な考え**	48.9	60.1	

* $p < .05$ ** $p < .01$ *** $p < .001$

ると、マーク及び旧来の記述式と、新傾向の記述式は、とくに新奇性において弁別されるようであった。この結果は、高校生の結果(宮本・倉元・田中, 2017)と同様であった。

解答に必要とされる資質・能力の判定においては、新傾向の記述式は、マークや旧来の記述式とは異なると捉え、一方、数学の新傾向の記述式に必要とされる資質・能力は、旧来の記述式と共通すると捉えている

ことが示唆された。また、全体的な布置をみると、国語の新傾向の記述式問題群は、国語の他の試験問題群よりも、数学の全試験問題群と近接していた。この結果は、高校生の結果(田中・宮本・倉元, 2018)と同様であった。

以上から、新傾向の記述式が持つ特異性の認知においては、高校教員と高校生で類似することが明らかになった。

ただし、高校教員の新傾向の記述式の印象評価や、解答に必要な資質能力の判定には、個別学力試験に対する日頃の検討よりも、改善の必要性といった問題意識が影響することが示唆された。改善の必要性を感じている者ほど、新傾向の記述式を肯定的に捉えるようであった。こうした一種の評価バイアスについては、留意する必要がある。

謝辞

本研究は、大学入学者選抜改革推進委託事業「個別学力試験『国語』が測定する資質・能力の分析・評価手法に関する研究(北海道大学(代表大学)、東北大学、九州大学、長崎大学、大学入試センター)の平成29年度の取り組みに基づくものである。

参考文献

- 中央教育審議会(2014) 新しい時代にふさわしい高大接続の実現に向けた高等学校教育、大学教育、大学入学者選抜の一体的改革について～すべての若者が夢や目標を芽吹かせ、未来に花開かせるために～(答申) 文部科学省 2014年12月22日
http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo0/toushin/_icsFiles/afeldfile/2015/01/14/1354191.pdf (2019年3月25日)
- 高大接続システム改革会議(2016). 高大接続システム改革会議「最終報告」 文部科学省 2016年3月31日
http://www.mext.go.jp/component/b_menu/shingi/toushin/_icsFiles/afeldfile/2016/06/02/1369232_01_2.pdf (2019年3月25日)
- 宮本友弘・倉元直樹・田中光晴(2017). 「問題の出題形式と測定する資質・能力の関係に関するテスト理論・測定学に基づく分析」 『日本テスト学会第15回大会発表論文抄録集』, 130 - 133.
- 田中光晴・宮本友弘・倉元直樹(2018). 「新共通テスト(イメージ例)が測定する資質・能力の分析—高校生対象のモニター調査から—」 『大学入試研究ジャーナル』 28, 1 - 6.
- 全国高等学校長協会編(2017). 『平成29年度版 全国高等学校一覧』 学事出版