

高大接続に資する学力評価測定の試み（国語）

Attempt of scholastic attainments evaluation measurement to contribute to connection of high school and university (Japanese)

西辻 正副
Masasuke Nishitsuji

1 研究の目的

本研究では、国語において、解答形式により試験のパフォーマンスに違いがあるのか？ということテーマに、試験によって測れる能力の解答形式による違いを実証的に調査することを通して、高大接続に資する学力評価測定について考察することを目的とした。

現行の大学入学者選抜における国語の試験の形式は、大学入試センター試験や私立大学の多くの入試問題で採り入れられている「選択形式」と、国公立大学を中心とした個別学力試験で多く採り入れられている「記述形式」の2つの形式が代表的である。そこで、国語においては^{注1)}、原則として、多肢選択形式（以下、「選択形式」という。）と記述形式とに分け、さらに、記述形式を自由記述形式（以下、単に「記述形式」という。）と条件を付した記述形式（一部に条件選択形式を含む。以下、「条件記述形式」という。）とに細分化し、合計3つの解答形式（以下、「形式」という。）の問題を作成し、調査することとした。

調査を通して、

- ① 3つの形式による難易度（得点）の差と、評価内容の違いを規定する要因は何か。
- ② 受験者の属性、学習履歴、問題に対する意識といったものが、形式による差にどの程度反映しているのか。を探ることを目的とした。なお、本稿では①を中心に報告を行う。

2 調査問題と調査の概要

調査問題は、国立教育政策研究所が実施した「平成14年度高等学校教育課程実施状況調査」の「国語Ⅰ」A冊子第2問（現代文（論理的な文章））の問題を許可を得てベースとして用い、各問いごとに選択形式（選択肢は5つ）、条件記述形式（漢字、語句等の一部の問いは、条件選択形式。）、記述形式の3種類の形式を作成した。そして、それぞれの形式が均等になるように配慮して、A、B、Cの3冊子（各50点満点）を作成した。

具体的には、次のような調査を行った。

- ① 能力が同等となるよう配慮した3つの集団に、1冊子ずつをそれぞれ解答させた。したがって、それぞれの集団の受験者は、内容は同じであるが形式が異なる問いを解答した。
- ② 3つの受験者集団の等質性を検証するために、受験者全員に共通問題として、共通一次試験（追試験）の過去問題（論理的な文章と古文の問題を抽出して1冊子（50点満点）を作成。以下、「共通問題冊子」という。）を解答させた。

③ 受験者の属性、学習履歴、問題に対する意識等の質問紙調査を実施した。

なお、調査の対象は、大学入学試験受験後あまり時間が経過していない時期（6月～11月）の大学1年次生（以下、「調査モニター」という。）とした。

平成23（2011）年度に試行的な調査を行い^{注2）}、その結果を分析して設問を見直し、平成24（2012）年度に本調査を行った。共通問題冊子は平成24（2012）年度のみ実施した。

各冊子の受験者が200名以上となることを目標に調査モニターを募り、844名の応募者のうち698名（参加率83%）が調査に参加し、各冊子とも目標以上となった。

3 結果と考察

(1) 冊子ごとの結果概要

冊子作成に当たっては、3つの形式を織り交ぜた3冊子の難易度が均等になるように配慮した。結果は、表1、図1に示すとおり、3冊子間で大きな差はみられなかった。

一部の問題で準正答を設定し、準正答と判断したとき、正答の得点を与える場合（配点×1）、正答の得点の半分の得点を与える場合（配点×0.5）、得点を与えない場合（配点×0）の3通りで採点した。

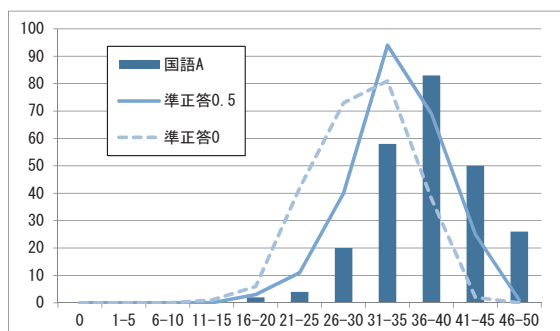
なお、本報告における得点及び正答率は、特に注記する場合以外は、準正答に正答の得点を与える場合（配点×1）で扱った。図1では棒グラフがそれに該当する。

【表1】

冊子	A冊子			B冊子			C冊子		
準正答(配点×)	1	0.5	0	1	0.5	0	1	0.5	0
平均点	37.7	33.7	29.7	34.4	30.4	26.4	31.4	29.6	27.7
標準偏差	6.1	5.1	5.1	8.0	7.9	8.4	9.0	7.9	7.2
最高点	49	46	44	50	47.5	45	50	49	48
最低点	18	15.5	13	11	8.5	6	7	7	7
人数	243			233			222		

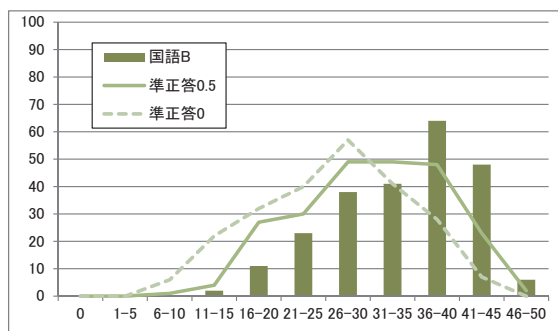
【図1】

(A冊子)

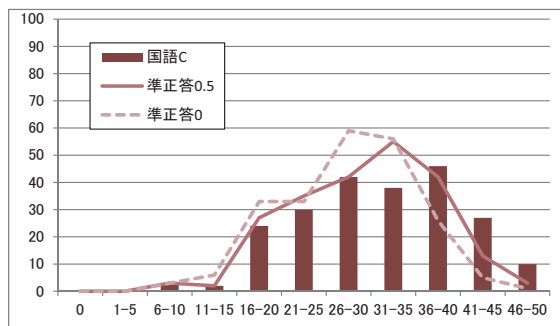


(B冊子)

(単位：縦軸（人）、横軸（点）)



(C冊子)



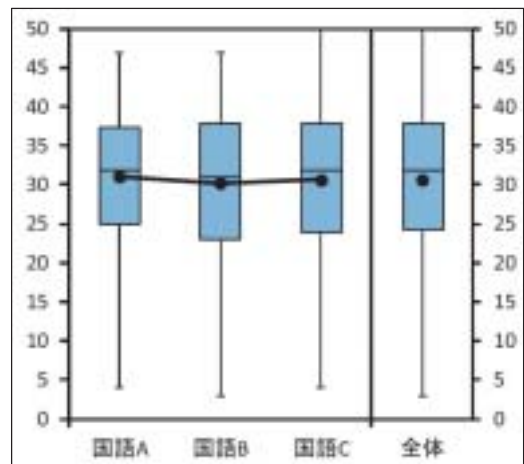
また、3つの受験者集団の等質性を検証するためのアンカーとして実施した共通問題冊子の結果をみると、次ページの表2、図2に示すとおりとなり、各冊子ごとの調査モニターの学力には、ほぼ差がないことが裏付けられたと判断した。

【表 2】

集 団	A 冊子	B 冊子	C 冊子
平 均 点	31.1	30.2	30.6
標 準 偏 差	9.1	9.6	9.8
最 高 点	47	47	50
最 低 点	4	3	4
人 数	243	233	222

【図 2】

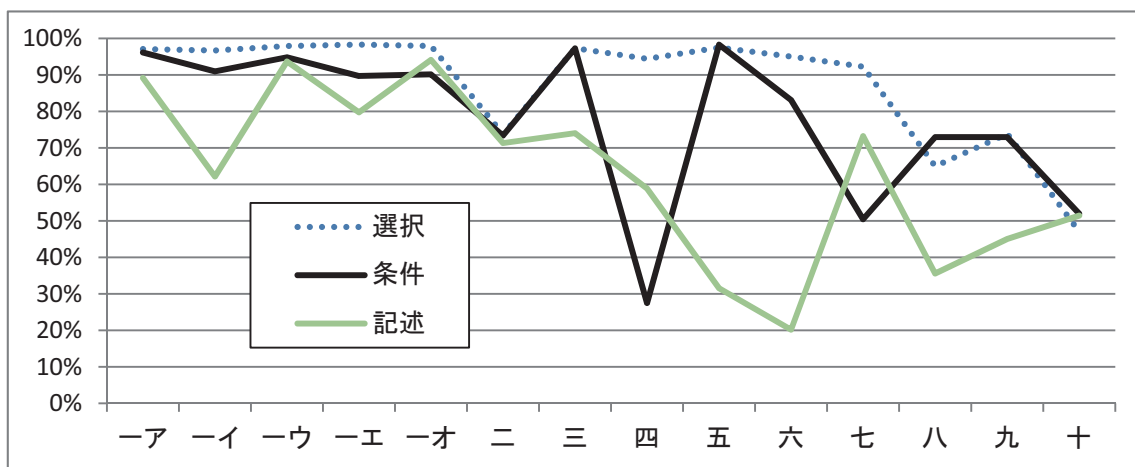
(単位：点)



(2) 設問ごとの形式別正答率比較

図3は、各設問の正答率（問い二、三、九は、小問をまとめて示している。）を形式別に示したものである。

【図 3】



予想されたとおり、選択形式が、条件記述形式及び記述形式よりもおおむね正答率が高かった。選択形式と記述形式とは正答率の差が大きく、測ることができる能力が異なっていると考えられる必要があると思われる。

条件記述形式の正答率は、おおむね記述形式の正答率を上回った。条件記述形式のうち、条件選択形式（問い一）や組合せ選択形式（問い二、三）は選択形式と同じかそれに近い正答率であった。一方、本文中からの抜き出しや、本文の内容を要約するなどの条件を付した問い四以降は、付された条件がヒントとなるか、制約となるかなど、条件の性格によって、選択形式の正答率に近いものもあれば、記述形式の正答率に近いものもあり、記述形式の正答率を上回ったものもあった。

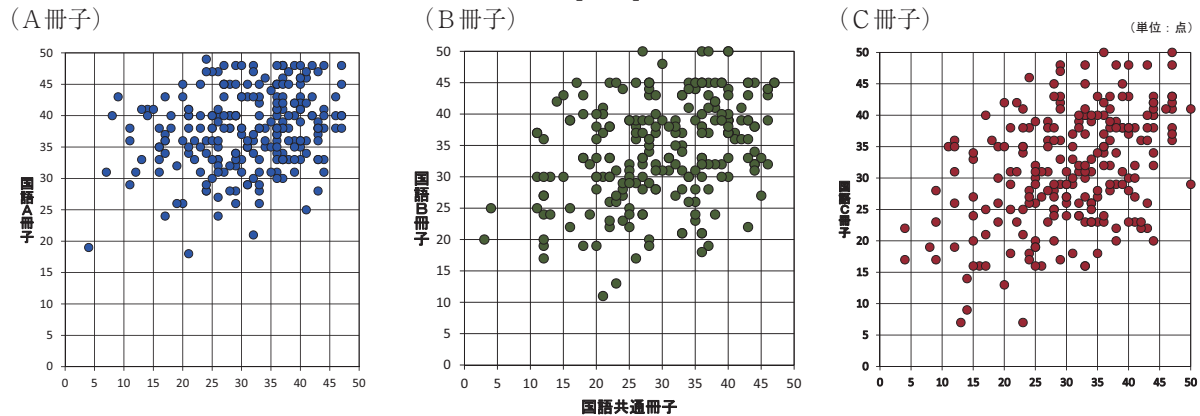
(3) 問題冊子と共通冊子の相関

共通冊子（全て選択形式）と、3つの形式が混在している3冊子の得点の相関は、次ページの表3、図4（正答の得点を与える場合（配点×1）のみ掲載。）のとおりであり、相関が低くなっている。同時に実施した数学が全て0.8近くあるのに比べて大きな違いがみられた。

【表3】

冊 子	A冊子			B冊子			C冊子		
	準正答 (配点×)	1	0.5	0	1	0.5	0	1	0.5
相関係数	0.239	0.267	0.245	0.360	0.376	0.360	0.428	0.384	0.330

【図4】



(4) 層ごとにみた各設問の形式別正答率比較

各冊子の設問について、得点の上位群、中位群、下位群の3層に分けて結果をみた。各群は、得点で調査モニターを四分位分割し、上位4分の1を上位群、下位4分の1を下位群、それ以外を中位群とした。

問い一 (図5)

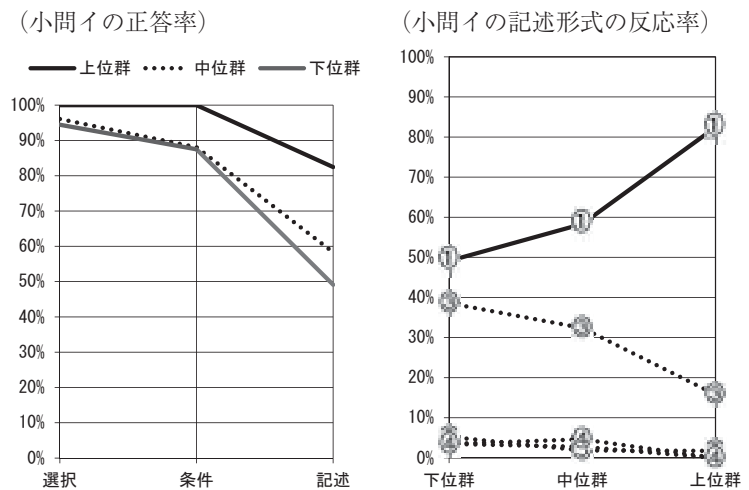
問い一は、文脈に即して正しい漢字を書くことができるかどうかをみる問題 (5問) である。

予測されたことではあるが、各層とも、おおむね選択形式 (選択肢は、漢字一字を示した。) の正答率が最も高く、条件選択形式 (選択肢は短文で示し、その中の当該熟語をカタカナで示した。) の正答率という順になった。選択形式では、各層間での正答率の差は小さく、条件選択形式、記述形式となるにしたがって、各層間での正答率の差が大きくなる傾向がみられた。

下位群は設問間の (設問の難易による) 正答率の差が各形式とも大きく、上位群は設問間の (設問の難易による) 正答率の差は各形式とも小さくなる傾向が見られた。

特色ある傾向を示したのは小問イ (規格化) (下線部が当該漢字。以下同じ。) である。この問いでは、記述形式の正答率の下位群及び中位群と、上位群との差が、他の設問に比べて大きかった。一方、下位群と中位群との差が、選択形式及び条件選択形式ではほとんどなかった。「規

【図5】



格」という語は、下位群及び中位群を中心に、一定数の学生にとって、なじみが少ない語であった可能性がある。この設問では、同音異義語の「企画」を、あえて選択形式及び条件選択形式の選択肢に含めていない。その結果、この2つの形式では、「規格」が「企画」よりなじみの度合いが低いことの影響が少なかったものの、記述形式でその影響が下位群と中位群とに顕著に表れ、「企画」と書いたものが下位群では38.6%、中位群では32.4%（図5の小問イの記述形式の反応率の③）^{注3)}にのぼり、上位群の正答率（図5の小問イの記述形式の反応率の①）との大きな差となったと考えられる。

下位群については、ア（子午線）、エ（精巧）にも、記述形式の正答率が他の2形式に比べて大きく下降する傾向がみられた。

なお、オ（余地）は、各層とも、記述形式が条件形式よりもわずかであるが上回っている。このことは、選択肢の熟語についてのなじみの度合いが影響していると思われる。

問二（図6）

問二は、文脈に即して正しい接続語を空欄（AからCの3箇所）に補充することができるかどうかをみる問題である。選択形式（3箇所の空欄ごとにそれぞれ適切なものを1つずつ選択。）、条件選択形式（当てはまる語の組合せを選択。以下、「組合せ選択形式」という。）、記述形式で出題した。

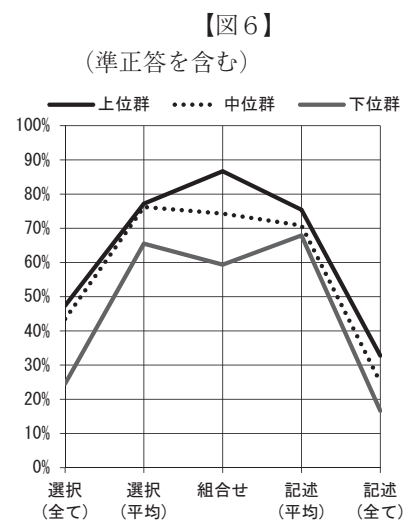
問題の本文中の一部分を空欄にして補充させる問題（以下、「空欄補充問題」という。）は、正答ありきの出題であり、特に選択形式の場合、学力を適切に測れないとして避ける傾向が強くみられるが、この調査では、解答形式比較の観点からあえて問いとして設定した（以下同じ）。

選択形式では、空欄Cについては逆接の語が入ることが文脈からみて比較的容易に分かることから正答率は高く、下位群から上位群へとその率も順に高くなっている。しかし、空欄A（結果を表す語）とB（話題転換を表す語）については、読み手の読みや考えではなく、筆者の用いた語を正答として選ぶというこの種の問題の性格から、中位群と上位群とに正答率の差がみられず、特にAについては上位群の正答率が中位群を下回り、この形式による出題の限界がみえている。3つの設問全てに正答した完答の正答率と、3つの設問の正答率の平均値とを層ごとにみると、下位群と、中位群及び上位群との間ではともに10ポイント以上の差がみられた。

組合せ選択形式では、逆接の語が入ることが文脈から比較的容易に分かる空欄Cが解答の鍵になり、各層とも多くの受験者は、選択肢をまずCに逆接の語が入っている2つに絞ることができたようだ。なお、上位群の正答率は、いずれの形式よりも高かった。また、各層間の差は、この形式が最も大きかった。

選択形式の正答率の平均値と組合せ選択形式とを比べると、上位群では、組合せ選択形式の正答率の方が10ポイント近く高くなっているが、下位群及び中位群では、逆に組合せ選択形式の正答率の方が低くなっている。

記述形式も、平易な空欄Cについては正答率が高かった。層ごとの差も選択形式とあまり変化はなかった。空欄A、Bについて上位群と下位群との差をみると、Aは選択形式とあまり変化はなかった。一方、Bについては、記述形式になると、層による差はほとんどなくなり、わずかの差であるが下位群の正答率が最も高かった。また、中位群及び上位群では「ところが」「それに対して」など話題転換（逆接や対比）を表す語を解答した



割合（正答）と、「その後」「やがて」など時間の経過を表す語（文脈からみて準正答と判断した）を解答した割合（準正答）とがほぼ拮抗している。

なお、記述形式において、3つの設問の正答率の平均値と、3つの設問全てに正答した完答の正答率とを層ごとにみた場合、正答のみでは各層間の差は小さかった。準正答（B、Cのみ）を加えると開きが出て、各層間の差が幾分はつきりとしている。

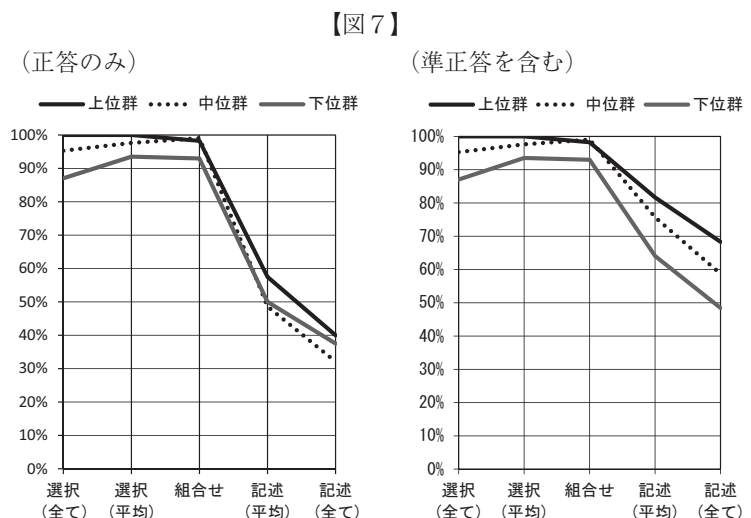
このようにみえてくると、空欄補充問題を出題する場合は、組合せ選択形式が、今回調査した3つの形式の中では、各層間の差をみるためには妥当なようにみえる。が、空欄補充問題では、「文脈からみて」正答となる語が複数あることから、この問いにおいても、選択形式、組合せ選択形式における選択肢は不確実さを避けるための配慮をしている（問い三も同じ）。したがって、正答率は高くなる傾向がみられる。なお、組合せ選択形式において、上位群はマス目のような形を示す選択肢群を論理的にみていく力が、下位群及び中位群よりもあるとみることができる。そこで、語彙の力そのものを純粹にみているかどうかについては、検討が必要であると考え。一方、記述形式においては、接続語を対象とした場合、文脈からみて、BやCのように正答を絞ることができず準正答を認める必要がある場合も多くなる。このようなことから、正答、準正答を合わせて正答として扱い、完答にしたり、配点比率を変えたりすること、選択肢にも準正答を含むことなどを検討する必要があると考える。

問い三（図7）

問い三も、問い二と同じく文脈に即した語の空欄補充問題である。問い二では接続語を対象としていたが、問い三では修飾語（ α 「あなたも絶対の実体」、 β 「まったく同じ」）を対象としている。類似した形式の設問であるこの2問の、選択形式と記述形式との正答率の差を比較すると、各層とも問い三の方が大きかった。

この問いは、選択形式、組合せ選択形式の正答率が高いにも関わらず、記述形式の正答率は高くない。修飾語の場合、選択肢があると、日常の言語生活における言語運用を背景に、その適否を語感から判断する傾向があり、ここでは、語彙の力というよりは言語感覚を測っている可能性がある。自ら適切な修飾語を考えることが求められる、記述形式の結果を（準正答も含めて）みてみることによって、上位群も含めて語彙の力が不足している傾向があり、その傾向は下位群がより深刻であることが初めてみえてくる。

記述形式における、2つの設問の正答率の平均値と、2つの設問の双方に正答した完答の正答率とを層ごとにみた場合、正答のみでは各層間の差は小さく、下位群が中位群を上回る逆転現象もみられるが、準正答を含めると、上位群と下位群との差は、平均値、完答ともに約20ポイントとなり、下位群と中位群との逆転現象もみられない。ここからも、語彙の力を測るためには、準正答を適切に活用することが求められると考える。



問い四（図8）

問い四も、語彙の力を測る空欄補充問題である。

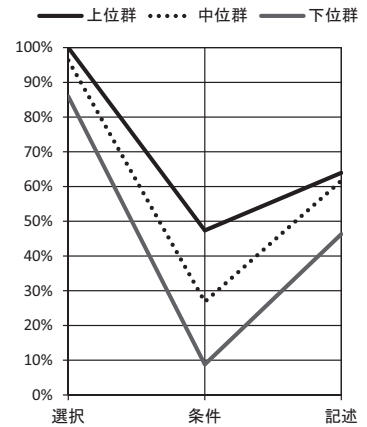
この問いは、「いいかんげんなものではない」というフレーズに続けるのにふさわしい語である「確固としたもの」を答えるものである。

選択形式では、中位群及び上位群は高い正答率を示しているが、下位群は、これらと10ポイント以上の差が出ている。

条件記述形式では補充する語を「漢語」に限定するという条件を付し、記述形式では条件を付さず和語等も可とした。その結果、「漢語」という条件が付いた条件記述形式における、上位群と下位群との正答率の差は約39ポイント、上位群と中位群との差も、条件を付さない記述形式では2ポイント余であるものが、約20ポイントまで広がり、層ごとの語彙の力の差が明確に表れている。

このことから、出題する際に、記述形式に適切な条件を付けることが、上位群と中位群の語彙の力の違いを的確に測る有効な手立てとなる場合があると考えられる。

【図8】
(準正答を含む)



問い五（図9・10）

問い五以降は、文章を読むことについての出題である。問い五から九については、選択形式と記述形式に、解答の文章の一部を空欄にしたものをあらかじめ示し、その空欄を補充する条件記述形式を加えた、3つの形式とした。

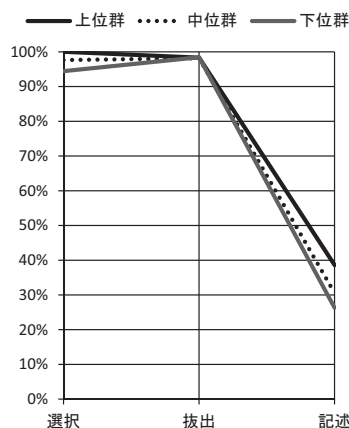
問い五では、筆者が強調していることの因果（相関）関係を的確に読み取る必要がある。

選択形式の正答率は高く、上位群は100%であった。選択肢の内容に曖昧さをなくしたことにより難易度が下がり、層ごとの差はほとんど生じていない。

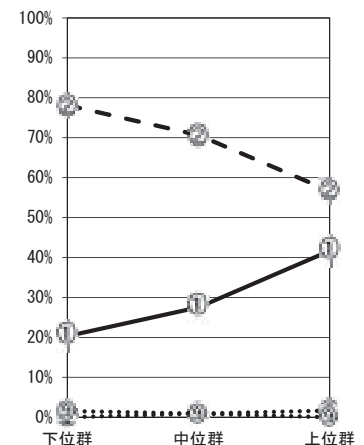
条件記述形式としては、空欄を本文中から抜き出して補充する形式（以下、「抜き出し形式」という。）を採用したが、選択形式の正答率との差や、層ごとの差はほとんど認められなかった。ただ、見かけとは異なり、層ごとの中身には大きな違いがみられた。各層の正答（図10の①）、準正答（図10の②）の割合を比較すると、正答については、上位群が下位群の約2倍（約20ポイント高い）となっている。逆に、準正答については、下位群が上位群よりも約20ポイント高くなっている。中位群は、上位群と下位群との中間のやや下位群寄りに位置している。このことは、下位群の読むことの質が、上位群のそれとは異なっていることを示しており、準正答の扱い方についての示唆を与えてくれる。

記述形式の正答率は他の2つの形式と比べて低く、各層とも4割以下となっている。これは、抜き出し形式では、「……により」と「原因」があらかじめ示され、「結果」を本文中から抜き出すだけであったのに対して、記述形式では、読み取った因果（相関）関係を、自らの力で適切

【図9】
(準正答を含む)



【図10】



に書くことが求められたことによると思われる。記述形式では因果（相関）のいずれかが欠けているものは準正答としなかった。因果（相関）関係が適切に書かれているもの（正答）、原因しか書かれていないもの（誤答）、結果しか書かれていないもの（誤答）の3者の割合は、各層とも大きな差はなかった。また、各層間の差も、正答のみでみた場合の抜き出し形式における差（図10の①）よりも縮小している。

このような結果から、因果（相関）関係を問う場合には、今回の抜き出し形式の問いのように、一方をあらかじめ与えて残りの一方を問うだけではなく、双方を抜き出させるなど、条件の示し方を変えて、記述形式と比較してみることも考えてよいと思う。

問い六（図11）

問い六では、文章の構成や展開に注意して、筆者の考えを捉える必要がある。

選択形式では、上位群は100%の正答率であったが、下位群はそれより10ポイント以上低かった。

条件記述形式では、空欄を補充する際に、「比べて」という語を用いることを条件とした。正答率は、選択形式と比べて各層とも10ポイント以上低下している。なお、上位群と中位群とは正答率の差がみられなかった。

この問いでは、記述形式に特徴が出ている。準正答を含めても、全体の正答率は約20%であり、上位群の正答率が中位群を10ポイント以上も下回るなどの逆転現象がみられた。

記述形式では、「大腸菌」と「人間」という説明の対象を2つとも明示する必要があるが、その一方あるいは両方を書いていないものが7割を超えていた。特に上位群には、文意を捉えているが、その内容を丁寧に説明する姿勢に欠けている者が多くいるように思われ、「大腸菌」は書いているものの、「人間」という対比の対象となるものは自明のものと勝手に思い込み、それを明示していない解答（図12の③）が6割を超えていた。このことは、読む能力はあるものの、目的や場に応じて、分かりきっていることでも過不足なく伝える必要があるという意識が弱いという傾向が、上位群に顕著に表れた結果と思われる。

問い七（図13）

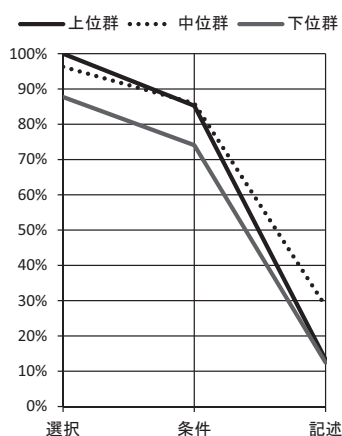
問い七では、文章の要旨を的確に捉える必要がある。

選択形式では、上位群の正答率は100%であり、各層間の差も明確にみられる。

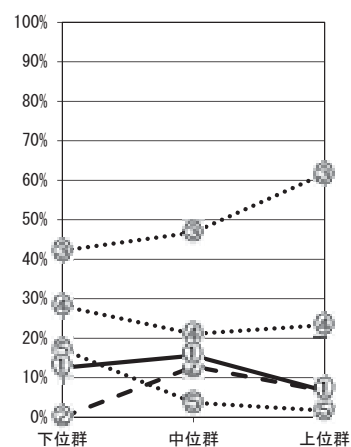
条件記述形式としての抜き出し形式では、選択形式と比べて各層とも約40ポイント低くなり、各層間の差もやや拡大している。

記述形式は、正答とした、文章の内容を一般化した記述をしているものだけでみると、各層とも抜き出し形式よりもさらに10ポイント以上（中位群は20ポイント超）低下している。ここでは、書かれていることを一般化できて

【図11】
（準正答を含む）



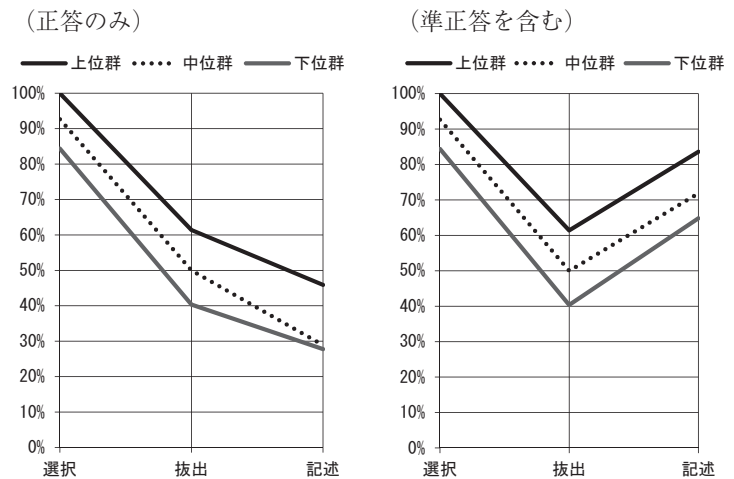
【図12】



いるかどうかで、上位群と、中位群及び下位群との読みの力の違いを的確に見分けることができると考えられる。文中の具体例を取り上げて解答しているものなども許容した準正答を含めると、各層とも、抜き出し形式よりも高い正答率となっているが、各層間の差は抜き出し形式とあまり変わっていない。

「読むこと」についての問いの中でも、この問いは、文章の要旨を捉えて解答することが求められるということで、難易度が高くなっている。そのため、抜き出し形式でも正答率が伸びていない。また、この難易度の高さによって、形式ごとの各層の正答率の変化が平行線に近くなっており、各形式とも、同質の読みの力を適切に測定し、各層間の差を明確に示しているということができると思われる。

【図13】

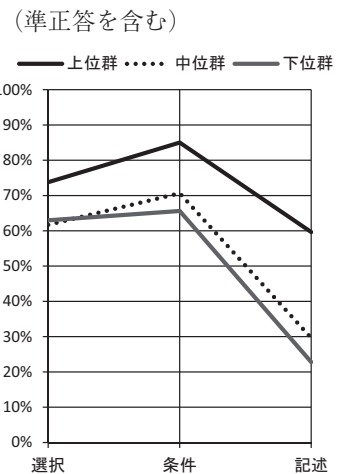


問い八（図14）

問い八では、「野蛮」という語にかぎっこを付けている、筆者の表現意図を捉える必要がある。この種の問いに受験者は慣れていないように思われ、選択形式においても、正答率は上位群で73.8%、下位群及び中位群では60%台の前半に止まっている。なお、正答とは全く正反対の内容の選択肢を、各層とも20%前後選択している。

条件記述形式では、「一見野蛮なように思えるが」という否定的な内容＋逆接の助詞に続けて空欄を補充することを条件として指定している。正答は、単なる肯定的な判断・評価の言葉のみではなく、具体的な内容を踏まえて一般化することまで求めており、肯定的な判断・評価の言葉のみものは準正答とした。この形式においては、「一見野蛮なように思えるが」と示したことが考えるヒントになり、準正答を含めた正答率は、中位群及び上位群では選択形式より高くなっている。正答と準正答との割合をみても、中位群及び上位群ではほぼ同率、下位群では準正答の割合が正答より約10ポイント高かった。この問いにおいても、準正答を適切に活用することで、各層の読みの質の違いをみる事が可能であると思われる。

【図14】



記述形式は、「野蛮」に対比する内容を適切に一般化したものを正答とし、文章の内容に依拠した具体的な記述のレベルにとどまっているものなどは準正答とした。その結果、正答のみの率は各層とも一桁台に止まった。準正答を含めると、上位群の正答率は59.6%となるのに対して、中位群は29.6%、下位群は22.8%と大きな差がついた。誤答をみると、中位群及び下位群では、内容が曖昧であったり、筆者の意図を間違えて捉えたりしている解答がともに50%を超えており、上位群と、中位群及び下位群との読みの力の差がはっきりとみえた。

これらの結果から、この問いは、難易度と形式の双方から、読みの力の高い層を見いだすのに適していると考え

られる。

問い九 (図15~17)

問い九では、文章の展開に注意して、筆者の考えの進め方を捉える必要がある。

選択形式は、下位群（正答率56.1%）と、中位群（正答率79.6%）及び上位群（正答率80.7%）との差が大きい（正答は図15の⑤）。この問いのような難易度の場合、選択形式において、読みの力の弱い層を見いだすことができると思われる。

条件記述形式としての空欄補充では、アで江戸時代の時計や時間について述べていることを文章内容からまとめ、イでまとめた内容を表すキーワードを文中から抜き出すという2段階の条件を課している。その結果、イの空欄の後の「入り込む余地がなくなった」がヒントとなり、イは各層とも90%台の正答率であったが、アの正答率は、「季節による違いへの対応」が何らかの形で書かれている準正答（図16の②）を含めても35.8%となった。正答のみ（図16の①）で見ると、上位群においても10%を切る結果であった。誤答をみると、中位群及び上位群では、対象は正しいがその内容を明確に書き切れていないもの（図16の③）が多かった。下位群では、それに加えて、題意に合わない内容（図16の⑤）を書いているものも多かった。

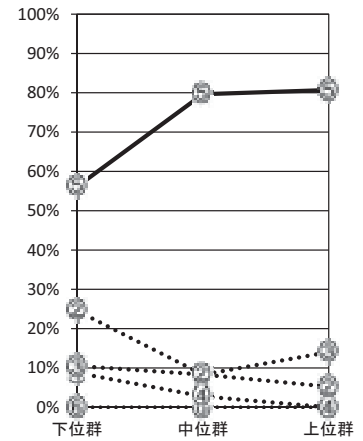
記述形式では、出現した状況（時間という絶対者の支配）と、その結果（人間の知恵を働かせる余地がなくなることを適切に書いているものを正答（図17の①）とし、準正答は設定しなかった。正答率は全体で45.1%となり、上位群は50%を超えている。誤答では、出現した状況だけを書いているもの（図17の②）が、下位群及び中位群でそれぞれ約40%おり、上位群との差が明確となっている。

問い十 (【図18】)

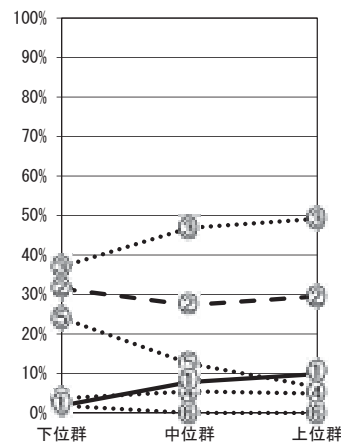
問い十では、文章のキーワードを的確に捉えるだけでなく、それをキーワードとした根拠を明確にする必要がある。

3つの形式のうち2つの形式では、あらかじめ示している4つの語からキーワードとして1つを選び、その語をキーワードと考えた理由を説明することを求めた。違いは、「根拠を明確にして」説明することを、傍点で強調しているか（選択+条件強調記述）、いないか（選択+記述）である。その結果、上位群では、傍点で条件を強調した場合の方が、しない場合よりも正答率が約17ポイント高くなり、中位群及び下位群との差も大きくなった。一方、中位群及び下位群では、条件強調の有無の影響はほとんどみられなかった。

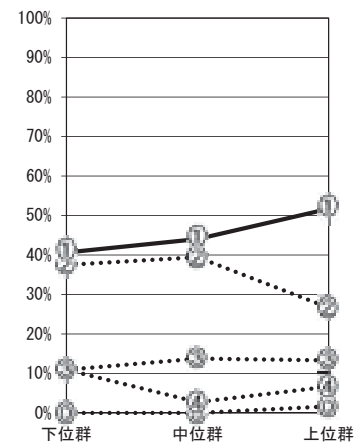
【図15】



【図16】



【図17】



残る1つの形式では、キーワードも含めて自由に記述させた。この形式でも「根拠を明確にして」説明することは傍点で強調した。選択+条件強調記述形式とは異なり、各層の正答率の差はほとんどみられなかった。上位群は、この形式の正答率が最も低かった。中位群及び下位群は、他の2形式よりも正答率が高かった。誤答としては、中位群及び下位群では、適切なキーワードを書けなかった割合が高く、上位群では、適切なキーワードを書けなかった割合と、キーワードは適切であるが根拠が適切に書けなかった割合が拮抗していた。

同じように条件強調していても、上位群にこのような正答率の差が出ているのには、与えられた語に根拠付けすることと、自分で探し出した語に根拠付けすることとに、異なる要素が影響している可能性がある。

さらに、この問いのようなオープンエンドの問題に受験者は慣れておらず、どのように答えてよいのか迷いがあったのではないかと考えられる。このような状況の下では、語を選ぶことから自由に記述させると、各層間の力の差を適切に測れないおそれがある。過渡期としては、選択形式を組み合わせたり、条件を効果的に強調したりするなどの工夫が当面求められると考える。

4 まとめ

全体の傾向として、形式による正答率の違いがみられた。これは国語特有のものと考えられ、特定の解答形式の正答率が高かったり、低かったりする特徴はみられず、問題の内容、設問の難易度、受験者集団の学力レベルに依存して変化するという傾向を示す数学^{注4)}とは異なる姿をみせている。

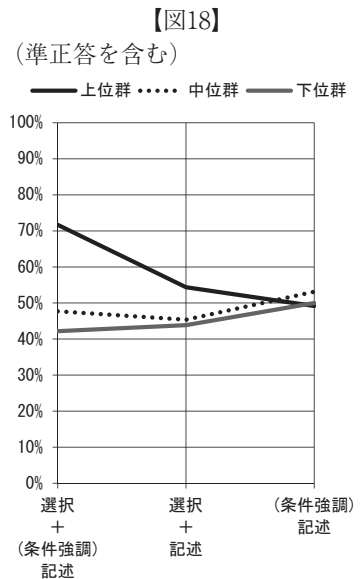
選択形式では、おおむね正答率が高くなり、各層間の差は小さくなる傾向がみられたが、これには、選択肢そのものが文章を読むことのヒントとなること、さらに、選択肢の中に必ず正答が1つあり、多くの場合、正答に揺るぎがないように誤答の選択肢を作るといふ、問題作成上の制約があることなども影響していると思われる。このことから、難易度が一定レベル以上でないと、層ごとの学力の質の違いを評価するには限界があると思われる。

条件記述形式（条件選択形式を含む。）では、その条件が、解答のヒントとなる場合（例えば問い八）と、かえって高いハードルとなる場合（例えば問い四）とがある。測りたい学力に応じた適切な条件設定が必要である。

問題の難易度や形式により、読みや語彙の力の高い層を見いだすのに適した条件や形式（例えば問い八の記述形式）や、読みや語彙の力の弱い層を見いだすのに適した条件や形式（例えば問い五の抜き出し形式、問い九の選択形式）があり、測りたい対象に適した難易度、条件、形式を活用することが求められる。

改めて選択形式、条件記述形式（一部、条件選択形式）、記述形式という3つの形式の結果を比較してみると、条件記述形式や記述形式によって広範な能力が測れることは言うまでもないが、はじめに正答ありきや、唯一絶対の正答にこだわるのではなく、正答に準正答を含めてゆるやかに考え、正答か、準正答かで異なる読みの質の差をみるなど、準正答の活用に一層注目することが大切である。そこで、選択形式の選択肢にも準正答を含め、それを選択した場合にも一定の点数を与えるなどの方式も考える必要がある。

なお、共通冊子と各冊子との相関が低かったことについては、今後、多様な角度から調査、分析を行う必要があると考える。



附記

本報告は、日本学術振興会科学研究費補助金基盤研究（A）「中等教育の多様化」に対応したコア学力の評価・測定を行うための技術的基盤の構築（平成21～25年度、課題番号21240069）（研究代表者は安野史子、西辻正副は研究分担者）の成果の一部である。結果データ等の詳細については、研究成果報告書「試験問題の形式とパフォーマンスに関する研究」に掲載している。

注

1) 本研究では、数学も同様の調査を行っている。なお、数学については、次の先行研究がある。

村上隆、三宅正武、藤村宣之、浪川幸彦、鈴木浩志、鈴木紀明、田栗正章、内田照久、安野史子（2008）「マークシート形式と記述形式による数学の「学力」（2）」『大学入試研究ジャーナル』、18, 163-170。

2) 2011年度の調査の概要については、数学の調査も併せて次の論考で報告している

安野史子、浪川幸彦、森田康夫、三宅正武、倉元直樹、林篤裕、木村拓也、宮埜寿夫、椎名久美子、荒井克弘、村上隆、西辻正副（2013）「解答形式とパフォーマンスに関する実証的研究」『大学入試研究ジャーナル』、23, 143-150。

3) 層ごとの反応率の結果をグラフで表したもの。グラフでは正答を実線、準正答を破線、誤答を点線で示し、マーカーの数字（①、②……）については選択形式が選択肢番号、それ以外の形式が解答類型番号を表している。以下、同じ。

4) 注2) に同じ。