

【原 著】

大学入学志願者の基礎的学力測定のための 「言語運用力」試作問題のモニター調査による分析 － 選択枝の変更が問題の特性に与える影響について －

荒井 清佳*
伊藤 圭*
椎名久美子*
宮埜 寿夫*
小牧研一郎**
桜井 裕仁*
田栗 正章***
安野 史子****

要 約

大学入学志願者の基礎的学力を測定するための「言語運用力」の問題が試作され、モニター調査が実施された。モニター調査では、複数の冊子が用意され、問題内容が同じであるが選択枝が異なる問題が実施された。本稿では、「言語運用力」の試作問題の中から、選択枝を変更して出題された6問を取り上げ、能力群ごとに示した各選択枝の選択率を比較することで、選択枝の変更が問題の特性に与える影響について分析する。分析の結果、選択枝の変更によって、正答率や識別力などの問題の特性が大きく変化すること、またその変化は受験者集団によっても異なることが示された。

キーワード：言語運用力，多枝選択式，項目分析，設問解答率分析図

1 はじめに

1.1 研究背景

大学進学率の上昇とAO入試や推薦入試による選抜の増加，大学入学志願者の高校での履修教科や科目の多様化（山村・荒牧，2002；山村，2010）に伴って，幅広い学力層が大学に入学する

ようになっている。また，平成12年11月の大学審議会の答申「大学入試の改善について」（大学審議会，2000）では，教科・科目にとらわれない基礎的，総合的な能力等の判定を目的とした総合的な問題に関する調査研究の必要性が指摘されている。

このような大学入学者選抜を取り巻く社会的要請のなか，大学入試センターでは，AO入試や推薦入試などを経て大学や短期大学に進学する受験

*独立行政法人大学入試センター研究開発部

**独立行政法人大学入試センター

***中央大学

****国立教育政策研究所

2014年3月14日 受理

者の基礎的な学力を診断・評価する「総合基礎」問題の試作（大学入試センター，2003，2004）や、医学部学士編入学者選抜を想定した教科・科目フリー型の総合試験の開発（柳井他，2006；林他，2010；伊藤他，2011），法科大学院における履修の前提として要求される判断力，思考力，分析力，表現力等の資質を測定する法科大学院適性試験の実施（椎名他，2007；杉澤他，2009）など，教科・科目別の試験とは異なる観点からの能力評価についての研究を行ってきた。平成23年度からは，これら先行研究の知見を踏まえて，従来の教科・科目型試験とは異なる観点から受験者の基礎的な能力や適性を評価する「新しい試験」の開発に関する新たな調査研究を進めている。

1.2 大学入学志願者の基礎的な学力測定のための「新しい試験」について

大学入試センター研究開発部で開発が進められている「新しい試験」とは，大学入学後の学修に必要な基礎的な能力を評価しようとする試験であり，受験者としてはAO入試や推薦入試などを経て大学や短期大学への入学を志願する幅広い学力層を想定している。この試験は「言語運用力」と「数理分析力」の二つの分野から構成されており，「言語運用力」では，課題遂行や問題解決を行う様々な場面で必要となる実践的な言語運用力を評価することを，「数理分析力」では，数理的な理解力，思考力，問題解決能力などを評価することを目指している（椎名他，2014）。

各分野の問題が試作され，2012年10月および12月には，高校2年生と大学・短大1年生を対象にモニター調査が行われた。このモニター調査では，複数の冊子が用意され，冊子間の比較ができるように計画された。冊子の構成や大問単位の予備的検討に関しては，「言語運用力」については椎名他（2014）および伊藤他（2014）に，「数理分析力」については桜井他（2014）に詳しく述べられている。

1.3 項目分析について

項目分析とは，予備的な調査等で得られた受験者の各問題項目への応答データを用いて，困難度や識別力等の項目の特性を分析することである。試験開発における重要な段階の一つであり，項目

分析の結果に基づいて利用する問題の選択が行われるほか，受験者の理解に関する情報ももたらされる（Hogan, T.P., 2007 繁 榊 他 訳, 2010；Downing, S.M. & Haladyna, T.M., 2006 池田監訳, 2008；Haladyna, T.M., 2004）。

本稿では，2012年12月のモニター調査で実施した「言語運用力」の問題のうち，冊子間で選択枝を変更して出題した問題に焦点を当て，選択枝の選択率に着目した分析を通して選択枝が問題の特性に与える影響について考察する。

2 言語運用力の問題と2012年度のモニター調査について

2.1 言語運用力の問題冊子

言語運用力の問題冊子は，素材文が日本語の大問が4問，素材文が英語の大問が4問の合計8問の大問で構成されている。いずれの問も多枝選択形式形式で，解答時間は40分である。問題冊子の全体構成については椎名他（2014）にまとめてある。日本語の問題は，一部の問題を除き，言語運用力を測定するために作られたオリジナル問題である。英語の問題は，一部がオリジナル問題であるほかは，過去に行われた大学入試センター試験の「英語リスニング」の問題を利用して作成したものである（伊藤他，2014）。次節で述べるように，言語運用力では4種類の問題冊子を作成したが，日本語の大問が4問，英語の大問が4問という構成はいずれも同じである。

2.2 2012年度のモニター調査

2012年度のモニター調査は2012年の10月と12月に実施された。10月の調査は予備的な調査として，AO入試や推薦入試での大学進学者が比較的多い高校（H1）の2年生を対象に小規模に行われた。その結果，正答率が低い問題や題意が分かりにくいという指摘が出た問題を中心に改良を図り，4種類の問題冊子（P，Q，R，S冊子）を作成した。P冊子は10月のモニター調査で用いたものと同じ冊子であり，Q，R冊子は10月の調査結果を踏まえて改良を図った冊子である。S冊子はできるだけ多くの試作問題の特性を把握

するため、初出の問題も含めた冊子である。

12月の調査は規模を拡大し、この4種類の問題冊子を用いて行われた。調査対象者は、一般入試での大学進学者が多い高校（H2, H3）の2年生とAO入試や推薦入試を経た入学者が多い大学と短期大学（U1～U3）の1年生である。なお、本稿では、この12月のモニター調査で実施されたP, Q, R冊子の結果を分析対象とする。S冊子は半分以上が初出の問題であるため、本稿での分析の対象外とする。

2.3 2012年度モニター調査における大問分析

日本語の問題に対する大問単位での分析は、R冊子の問題を対象に椎名他（2014）で行われている。そこでは高校生／大学生別に大問得点率分析図を描き、両者の折れ線を比較している。その結果、高校生では層によっては識別のみられないところがあるような大問であっても（第1問, 第4問, 第7問）、大学生では下位層から上位層まで順調に得点率が上がり、高い識別力を持つことが示された。また、第5問は高校生、大学生ともに正答率が低く、アンケート結果とも合わせると設問文を工夫するなどの必要があることが分かった。

英語の問題に対する分析は、伊藤他（2014）が設問正答率分析図を用いた分析を行っている。その結果、英語の各問題についても高校生に比べて大学生において識別力が高いことが示された。また、選択枝を変更して出題した3問（第2問の問2および第8問の二つの設問）についての分析も行っており、その結果、問題内容が同一でも選択枝を変化させることによって問題項目の困難度を操作できることが示された。

3 分析

3.1 分析対象

本稿では、2012年12月のモニター調査で実施されたP, Q, R冊子の結果を用いる。表1に、対象となる問題冊子と学校別の受験者数を示す。

P, Q, R冊子に含まれる問題のうち、本稿で分析対象とするのは問題内容が同じだが選択枝が異なる問題である。対象となる問題は6問あり^{注1)}、表2にその問題IDを示す。問題IDのアルファ

表1 問題冊子と学校別の受験者数

実施校	P冊子	Q冊子	R冊子
H2 高校	76	39	36
H3 高校	79	79	40
高校小計	155	118	76
U1 短大		34	33
U2 大学		39	30
U3 大学		38	35
大学小計		111	98
計	155	229	174

ベットは問題内容を、数字はその選択枝のバリエーションを示しており、同じアルファベットで数字が異なる場合は、問題内容は同じであるが選択枝の内容が異なっている問題であることを示す。英語の問題は、伊藤他（2014）でも分析対象となっている問題であり、伊藤他（2014）と同じ問題IDを付した。

表2 問題ID

問題番号	解答番号	日/英	P冊子	Q冊子	R冊子
第1問	問1	[1]	日本語	m1	m2
第2問	問2	[4]	英語	b1	b2
第5問		[10]	日本語	p1	p2
第6問	問3	[16]	日本語	q1	q2
第8問		[17]	英語	j1	j2 [17] j3 [17]
第8問		[18]	英語	j1	j2 [18] j3 [18]

3.2 分析方法

以下では設問解答率分析図を描いて分析を行う。設問解答率分析図とは、冊子の得点の順に受験者集団をほぼ等しい人数のいくつかの群に分け、各選択枝の選択率を受験者群ごとに求めて折れ線をつないだものである。正答選択枝の折れ線だけでなく、誤答選択枝の折れ線も描くことで、誤答選択枝の役割を分析することが可能となる。折れ線が高い位置にあるほど、その受験者群によって多く選択された選択枝となる。また、正答選択枝の折れ線については、傾きが大きければ、その傾きの大きい群間の違いを識別する力が大きいことを意味する。本稿では受験者集団を5群に分け、上位から順に、上位群（H）、上中位群（HM）、中位群（M）、下中位群（LM）、下位群（L）と

した。また、すべての選択枝を図示すると煩雑になってしまうため、全体での選択率が5%未満の選択枝は省略した。

受験者集団の人数が少ない場合には、わずかな差が選択率に大きな違いを与えてしまう。2012年12月のモニター調査のデータの場合、R冊子はH2高校とH3高校を合わせても76人であり、5群に分けて分析するには人数が少ない。そこで、本稿では受験者集団をまとめることにした。問題m, b, p, qについては、Q冊子とR冊子で全く同じ問題である。そこで、Q冊子、R冊子を受けた高校生、大学生の四つの受験者集団をそれぞれ5群に分け、高校生の対応する群同士、大学生の対応する群同士を合体させて分析し、結果を高校生(194人)／大学生(209人)別に示した。問題jについては、Q冊子とR冊子で異なる選択枝を提示している。そこで、高校生と大学生のデータを合わせ、Q冊子を受けた229人とR冊子を受けた174人をそれぞれ5群に分けて分析した結果を示した。表3に、各問題の設問解答率分析図を作成する際に用いた集団と人数を示す。以下、「高校生」はH2高校およびH3高校の受験者を合わせた集団を、「大学生」はU1短大、U2大学およびU3大学の受験者を合わせた集団を指すこととする。

3.3 問題 m の分析

図1に問題mを示す。問題m1はP冊子で出題したもので選択枝は五つである。問題m2はQ、R冊子で出題したもので、選択枝を四つに減らしたものである。問題mの設問解答率分析図を図2に示す。図2(a)は問題m1について、図2(b)

(c)は問題m2についての設問解答率分析図であり、(b)は高校生、(c)は大学生の結果である。図2の凡例で示す数字は選択枝番号であり、正答選択枝は①で、太い実線で表している。正答率は、図2の(a)(b)(c)の順に、0.57, 0.57, 0.40である。

図2(a)と図2(b)はどちらも高校生についての分析結果である。問題m1では五つの選択枝のうち、図に現れた選択枝は二つだけであり、残りの三つの選択枝は選択率が5%に満たなかった。それに対して問題m2では四つの選択枝とも図に現れ、すべての誤答選択枝が5%以上選択されている。誤答選択枝の中で選択率が最も高いのは、問題m1での選択枝⑤(問題m2での選択枝④)である。この選択枝の後半には「AとBは友達である」という、設問で与えられた文章「AとBは友達ではない」の正反対の表現が含まれているため、「間違っているもの」を文章の意味を考えずに表面的に選ぼうとすると選びやすい選択枝であると考えられる。また、この選択枝の前半の「AともBとも友達でない」が設問の文章と見た目と同じであるため、間違っているものを選ぶという設問に注意しない場合にも選びやすい選択枝であると考えられる。問題m1, m2のどちらの問題でもこの選択枝を下位層ほど選択しており、問題mの識別性に寄与していることが分かる。

図2(b)と(c)は問題m2の高校生と大学生での結果である。両者を比べると似たような傾向にはあるが、高校ではLM群で選択枝①と④が交差しているのに対し、大学ではM群・HM群で交差しており、H群のみが正しく答えられていることが分かる。「新しい試験」が想定する受験

表3 各問題の設問解答率分析図に用いた集団と人数

問題ID	m1, b1, p1, q1	m2, b2, p2, q2 (H)	m2, b2, p2, q2 (U)
集団と人数	高校生 (P冊子) 155人	高校生 (Q冊子 + R冊子) 194人	大学生 (Q冊子 + R冊子) 209人
問題ID	j2	j3	
集団と人数	高校生 + 大学生 (Q冊子) 229人	高校生 + 大学生 (R冊子) 174人	

者に近いのは今回のモニター調査では大学生であり(椎名他, 2014), その大学生においては上位層とそれ以外の層とを識別できている。

3.4 問題 b の分析

図 3 に問題 b を示す。問題 b1 は P 冊子で出題したもので選択枝は四つである。問題 b2 は Q, R 冊子で出題したもので、選択枝を六つに増やし、さらに誤答選択枝のすべてを設問中に登場する数字から起こりうる誤りパターンで構成した問題である。問題 b の設問解答率分析図を図 4 に示す。図 4 (a) は問題 b1 について、図 4 (b) (c) は問題 b2 についての設問解答率分析図であり、(b) は高校生、(c) は大学生の結果である。図 4 の凡例では選択枝の内容を示しており、正答は 44 ドルで太い実線で表している。正答率は図 4 の (a) (b) (c) の順に、0.28, 0.16, 0.06 である。

問題 b では、save と free という単語の意味を正しく理解し、会話の文脈にそって解釈できているかどうかの問題のポイントとなると考えられる。その場合、両方を正しく解釈できれば正答の 44 ドルとなるが、save を解釈できないと 55 ドル、free を解釈できないと 56 ドル、両方とも解釈できないと 70 ドルになる。

図 4 (a) より、問題 b1 では正答選択枝の折れ線を見ると、正答率は高くはないものの、LM 群以下と M 群以上とを識別できていることが分かる。誤答選択枝の折れ線を見ると、いずれの群でも 56 ドルの選択率が高く、free の意味を解釈できない受験者が一定数いることが分かる。下位層では 70 ドルの選択率も高い。しかし問題 b1 では「55 ドル」の選択枝がないため、free は解釈したが save は解釈できていない受験者(55 ドルを選ぶと予想される受験者)の割合は分からない。それ

【問題 m1】
 次の文から考えられる状況として、間違っているものを、下の①～⑤のうちから一つ選べ。
 「AさんとBさんは友達ではありません」

① Aは私ともBとも友達であるが、私はBとは友達ではない。
 ② 私はAともBとも友達であるが、AはBとは友達ではない。
 ③ 私はAとは友達であるが、Bは私ともAとも友達ではない。
 ④ Aは私ともBとも友達でないが、私とBは友達である。
 ⑤ 私はAともBとも友達でないが、AとBは友達である。

【問題 m2】 (設問は m1 と同じ)

① Aは私ともBとも友達であるが、私はBとは友達ではない。
 ② 私はAともBとも友達であるが、AはBとは友達ではない。
 ③ Aは私ともBとも友達でないが、私とBは友達である。
 ④ 私はAともBとも友達でないが、AとBは友達である。

図 1 問題 m

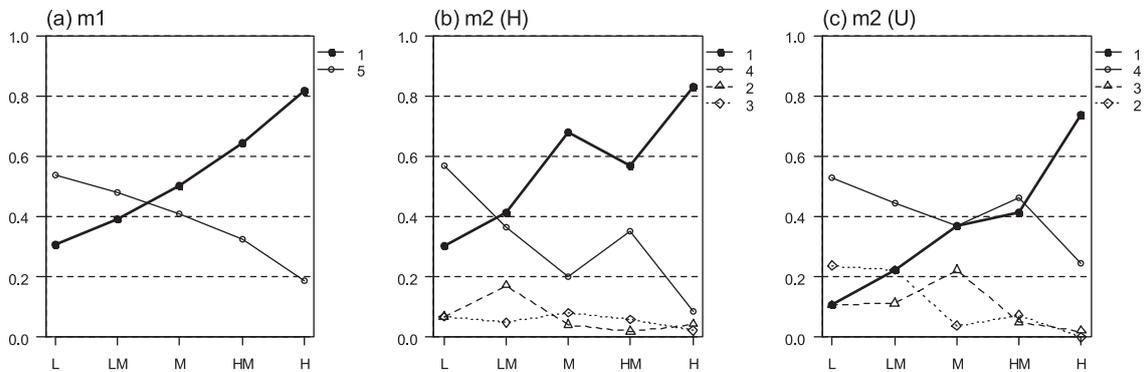


図 2 問題 m の設問解答率分析図 ((a) 問題 m1, (b) 問題 m2 (高校生), (c) 問題 m2 (大学生))

【問題 b1】
 次の英文は、小さな美術館の前での会話である。教師 3 人、生徒 11 人のグループの入場料は何ドルか。正しいものを、下の①～④のうちから一つ選べ。
 M: The sign says admission is five dollars each.
 W: But since we're a group, we can save a dollar on each ticket.
 M: We have eleven students and teachers are free.
 W: Sounds good.
 ① 40 ドル ② 44 ドル ③ 56 ドル ④ 70 ドル

【問題 b2】 (設問は b1 と同じ)
 ① 44 ドル ② 52 ドル ③ 55 ドル ④ 56 ドル ⑤ 65 ドル ⑥ 70 ドル

図 3 問題 b

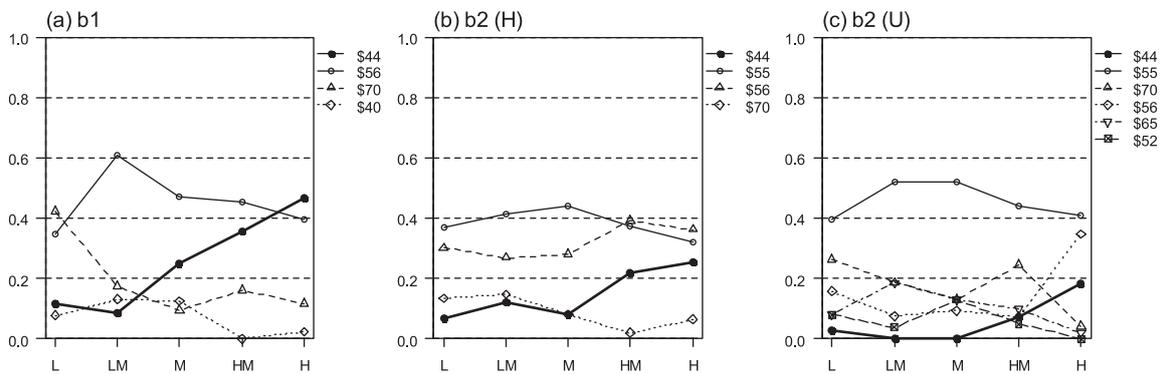


図 4 問題 b の設問解答率分析図 ((a) 問題 b1, (b) 問題 b2 (高校生), (c) 問題 b2 (大学生))

に対して、図 4 (b) の問題 b2 を見ると、いずれの群でも 55 ドルと 56 ドルの選択率が高いことが分かる。問題 b1 に比して 70 ドルの選択率は低い。ここから、高校生は save あるいは free のどちらかの単語を解釈し損ねて間違えてしまっている人が多いことが分かる。また図 4 (c) より、大学生では 55 ドルの選択率が高く、56 ドルの選択率は低い。ここから大学生は、save は解釈できない者が半分近くいるものの、多くは free の意味は解釈できるということが推察される^{注2)}。問題 b2 は選択枝の変更によって、誤答選択枝の分析を通じて受験者がどこで間違えているのかが分かる問題へと変更されたと言える。

3.5 問題 p の分析

図 5 に問題 p を示す。問題 p1 は P 冊子で出題したものである。問題 p2 は Q, R 冊子で出題したもので、選択枝①の内容を変更したものである。問題 p の設問解答率分析図を図 6 に示す。図 6 (a) は問題 p1 について、図 6 (b) (c) は問題 p2 に

ついての設問解答率分析図であり、(b) は高校生、(c) は大学生の結果である。図 6 の凡例の数字は選択枝番号である。正答選択枝は③で、太い実線で表している。正答率は、図 6 の (a) (b) (c) の順に、0.39, 0.40, 0.23 である。

図 6 (a) より問題 p1 では正答率がよくはないものの、LM 群以下と M 群以上とを識別できていることが分かる。誤答選択枝を見ると、選択枝②は上位の群になるにしたがって選ばれなくなっていくが、選択枝④は群によらず 2～4 割が選択している。また選択枝①はほとんど選ばれていない。問題 p2 はこの選択枝①を変更した問題である。しかし、図 6 (b) から分かるように、問題 p2 でも選択枝①はほとんど選ばれず、正答選択枝の折れ線の形状も L 群を除いて問題 p1 と似たような形となった。問題を見ると、問題 p1 の選択枝①は誤りであると分かりやすい選択枝であった。それに比べて問題 p2 の選択枝①は分かりにくくなっているが、他の選択枝よりは誤答と分かりやすい。ここから言えることは、選択枝を変更

【問題 p1】次に示す推論の中に、正しい推論が一つだけある。正しい推論を、次の①～④のうちから一つ選べ。

①ある日本人は陽気である。
ある日本人は太っている。
ゆえに、陽気な人は太っている。

②火星人は金星人ではない。
地球人は火星人ではない。
ゆえに、地球人は金星人ではない。

③カツ丼を食べれば試験に合格する。
彼女は試験に合格しなかった。
ゆえに、彼女はカツ丼を食べなかった。

④国会議員は政治家である。
私の妻は政治家である。
ゆえに、私の妻は国会議員である。

【問題 p2】（設問と選択枝②～④は p1 と同じ）

①詩人は恋する者ではない。
彼は詩人である。
ゆえに、彼は恋する者である。

図 5 問題 p

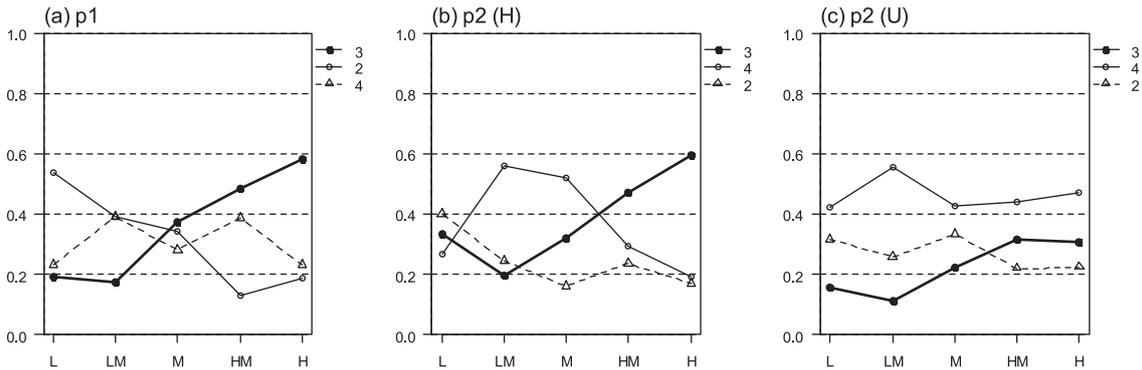


図 6 問題 p の設問解答率分析図 ((a) 問題 p1, (b) 問題 p2 (高校生), (c) 問題 p2 (大学生))

したとしても、同様に誤りであることが分かりやすい選択枝なのであれば、選択枝の変更はそのほかの選択枝の選択に影響をほとんど与えないということである。

また、図 6 (c) を見ると、大学生では H 群であっても、正答選択枝よりも誤答選択枝④の選択率が高いことが分かる。しかし、選択枝①は高校生と同様に選ばれていない。問題 p は大学生での正答率が 0.23 と低く、大学生には難しい問題であった。難しい問題であっても選択枝①のように分かりやすい誤りは選択されていないことから、難しいからといって受験者が全くランダムに選択枝を選ぶということはないということが分かる。

3.6 問題 q の分析

図 7 に問題 q を示す。問題 q1 は P 冊子で出題したもので選択枝が六つである。問題 q2 は Q, R 冊子で出題したもので、選択枝の②と⑤を削除して四つにしたものである。問題 q の設問解答率分析図を図 8 に示す。図 8 (a) は問題 q1 について、図 8 (b) (c) は問題 q2 についての設問解答率分析図であり、(b) は高校生、(c) は大学生の結果である。図 8 の凡例の数字は選択枝番号である。正答選択枝は①で、太い実線で表している。図 8 (c) の *は無効な回答^{注3)}である。正答率は、図 8 の (a) (b) (c) の順に、0.41, 0.35, 0.26 である。

【問題 q1】

(問題文及び問 1, 問 2 は省略)

問 3 地中を伝わる地震波には、振動の向きと伝わる速度が異なる S 波と P 波という 2 種類の波があり、これまでの観測から、震源地から見て地球の反対側のある範囲には、P 波は伝わるが、S 波が伝わらない「影」の領域があることが知られている。この現象が起きる理由を説明するために、本文に書かれていることからのほかに必要となる知識を記した文として最も適当なものを、次の①～⑥のうちから一つ選べ。

- ① P 波は固体中も液体中も伝わり、S 波は固体中は伝わるが液体中は伝わらない。
- ② S 波は固体中も液体中も伝わり、P 波は液体中は伝わるが固体中は伝わらない。
- ③ P 波は固体中も液体中も伝わり、S 波は液体中は伝わるが固体中は伝わらない。
- ④ P 波は金属中も岩石中も伝わり、S 波は金属中は伝わるが岩石中は伝わらない。
- ⑤ S 波は金属中も岩石中も伝わり、P 波は金属中は伝わるが岩石中は伝わらない。
- ⑥ S 波は金属中も岩石中も伝わり、P 波は岩石中は伝わるが金属中は伝わらない。

【問題 q2】(設問は q1 と同じ)

- ① P 波は固体中も液体中も伝わり、S 波は固体中は伝わるが液体中は伝わらない。
- ② P 波は固体中も液体中も伝わり、S 波は液体中は伝わるが固体中は伝わらない。
- ③ P 波は金属中も岩石中も伝わり、S 波は金属中は伝わるが岩石中は伝わらない。
- ④ S 波は金属中も岩石中も伝わり、P 波は岩石中は伝わるが金属中は伝わらない。

図 7 問題 q

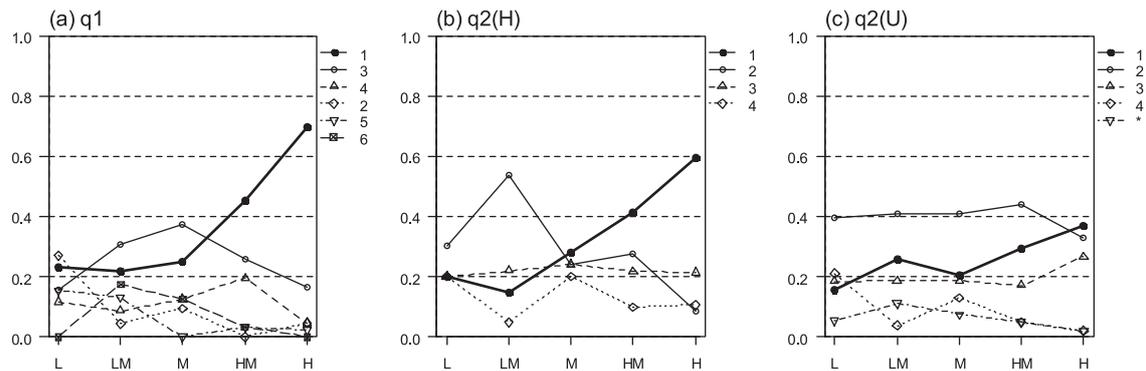


図 8 問題 q の設問解答率分析図 ((a) 問題 q1, (b) 問題 q2 (高校生), (c) 問題 q2 (大学生))

問題 q は、地球の内部構造についての説明文を読んだ上で問 3 での間に答える問題であり、地震波に関する知識が無くても答えられるようになっている。図 8 より、問題 q1 と問題 q2 のどちらも選択枝①, ③, ④ (問題 q2 ではそれぞれ選択枝①, ②, ③) の選択率が高い。これは問 3 では「S 波が伝わらない理由」を問われているため、「S 波は…伝わらない」という言葉が入っている選択枝を選んだためと考えられる。図 8 (a) と (b) を比べると、高校生では正答選択枝の折れ線は両者とも大きな違いはなく、問題 q1 では六つの選択枝のすべてが、問題 q2 では四つの選択枝のすべてが図に表れており選択枝として機能している。ここから、問題 q では、誤答選択枝のいずれもが正解が分からない者には魅力的で選

択しようと思わせるような選択枝になっているということが分かる。

図 8 の (b) と (c) を比べると、大学生ではいずれの群でも誤答選択枝②を選択する者が多く、また、無効な回答が多い。大学生での無効な回答の割合は 5.7% であり^{注 4)}、全部で 18 ある設問中で最も高かった。問題 q は問題冊子の後ろの方にあること (18 設問中 16 番目) に加え、文章を読ませた上でさらに考えさせる問題であったので、時間不足によりきちんと答えられなかったのかもしれない。

3.7 問題 j の分析

図 9 に問題 j2 及び j3 を示す。問題 j1 は、2012 年 10 月に実施した予備的な調査での正答率が極

端に低かったため、Q、R 冊子では問題内容を簡単にし、さらに選択枝の組合せを Q 冊子と R 冊子で変更した。本稿では、Q 冊子に出題された問題 j2 と R 冊子に出題された j3 の比較を行う。解

答番号 17 と 18 の 2 問をそれぞれ j [17] , j [18] と解答番号を付けて区別する。問題 j [17] の設問解答率分析図を図 10 に、問題 j [18] の設問解答率分析図を図 11 に示す。図 10 及び図 11 の(a)

【問題 j2】
 雄鹿 (stag) は、9 月の終わり頃から 10 月の初め頃にかけて、雌鹿の群れの所有をめぐって争うことが知られている。次の文章を読み、図中の空欄 17・18 に入る数をそれぞれ下の①～⑨のうちから一つずつ選べ。ただし、同じものを繰り返し選んでもよい。

Figure 1 shows 50 contests between two stags, which approached to within 100 meters of each other. Of these only 14 led to escalated fights, and in 8 of these cases escalation was preceded by a roaring contest.

```

    graph LR
      A["Approach (50)"] --> B["Roar contest (33)"]
      A --> C["No roar contest (17)"]
      B --> D["Fight (17)"]
      B --> E["One stag withdraws (3)*"]
      C --> F["Fight (?)"]
      C --> G["One stag withdraws (8)*"]
    
```

Figure 1

※ withdraw (引き下がる)

① 4 ② 5 ③ 6 ④ 7 ⑤ 8 ⑥ 9 ⑦ 10 ⑧ 11 ⑨ 12

【問題 j3】 (設問は j2 と同じ)

① 5 ② 6 ③ 8 ④ 9 ⑤ 11 ⑥ 14 ⑦ 17 ⑧ 25 ⑨ 27

図 9 問題 j

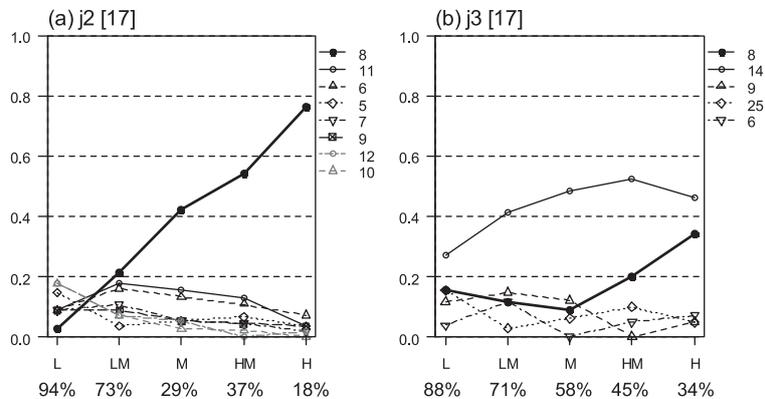


図 10 問題 j [17] の設問解答率分析図 ((a) 問題 j2 [17], (b) 問題 j3 [17] ; 群名の下の数値は、その群における大学生の割合 (%) である)

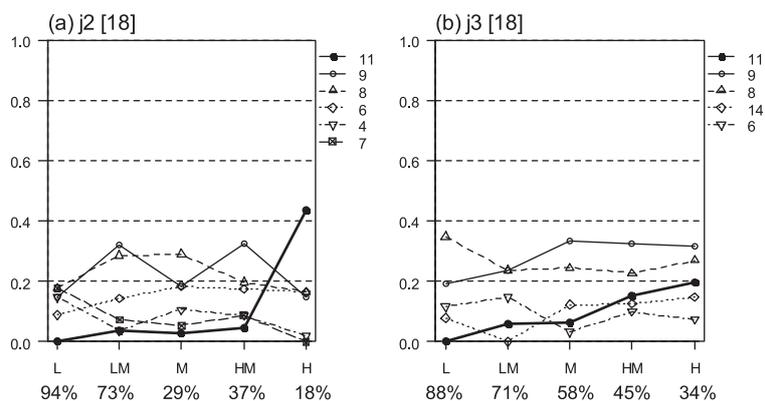


図 11 問題 j [18] の設問解答率分析図 ((a) 問題 j2 [18], (b) 問題 j3 [18])
; 群名の下の数値は, その群における大学生の割合 (%) である)

は問題 j2 の, (b) は問題 j3 の分析図である. これまでと異なり, 問題 j では高校生と大学生のデータをまとめて分析しているため, 各群における大学生の割合 (%) を図中に示した. 図 10, 図 11 とともに凡例の数字は選択枝番号ではなく, 選択枝の内容 (数値) を示している. 問題 j [17] の正答は 8, 問題 j [18] の正答は 11 であり, いずれも太い実線で表している. 正答率は, 図 10 の (a) (b) の順に 0.42, 0.19, 図 11 の (a) (b) の順に 0.13, 0.10 である.

問題 j の難しいところは, 英文が学術的な文章であり, 読み慣れない内容であることに加え, 英文から読み取った情報をもとに図中の空欄に入る数字を計算して求めなければならない点である. 問題 j [17] は英文中の「8」を読み取れば答えられるが, 問題 j [18] はその「8」を利用して図中で伏せられている数値を求め, さらにその数値を用いて計算する必要がある. 問題 j2 と問題 j3 の違いは, 前者は選択枝に 4 ~ 12 の数を順に並べただけであるのに対し, 後者は選択枝に英文中や図中に登場する数から計算できる数を並べているところである.

図 10 より, 問題 j2 [17] では H 群の正答率が 0.8 近くあり識別力も高いのに対し, 問題 j3 [17] では H 群でも正答率が 0.4 に届かず, いずれの群でも誤答選択枝である「14」の選択率が正答選択枝の「8」を上回っている. 「14」は英文中に登場する数字の一つで, かつ, fight という単語と同じ文の中に含まれるため, 受験者が選びやすい誤答選択枝になっていると推察される. 問題 j2 [17]

を受けた受験者と問題 j3 [17] を受けた受験者は, ほぼ同等な集団と見なせるにもかかわらず, 正答率で大きな差が生じている. その理由としては問題 j2 [17] の選択枝には「14」が存在しないという点が考えられる. 問題 j3 [17] で多く選ばれた誤答選択枝「14」が, 英文中に登場するという理由で選択された可能性が高いことから, 問題 j2 [17] においても, 正しく読めていないにもかかわらず, 英文中に登場するという理由で「8」を選択して偶然正解した者もいると考えられる.

また伊藤他 (2014) でも述べられているが, 問題 j [17] では, 選択枝番号と答えたい数字とを混同した人がいる可能性もある. 問題 j2 [17] で 2 番目に多く選択された「11」は選択枝⑧であり, 問題 j3 [17] で 4 番目に多く選択された「25」も選択枝⑧である. 選択枝の内容が一桁の数字である場合には選択枝番号と混同しやすいので, 例えば選択枝に「8匹」のように単位を付けるなどの何らかの工夫をすべきであったらう.

続いて図 11 を見ると, 問題 j2 [18] では H 群の正答率が 0.4 以上と高いのに対し, 問題 j3 [18] では正答率も低く識別も低い. ただし, H 群に占める大学生の割合を見ると, 問題 j2 では 18%, 問題 j3 では 34% である. 今回のモニター調査では全体的に大学生の方が高校生よりも正答率が低いので, 問題 j2 の H 群の大学生の割合が問題 j3 のそれと同程度であったとしたら, H 群の正答率は下がると考えられ, 正答選択枝の折れ線の形状の違いは H 群に占める大学生の割合の違いによる可能性がある. また問題 j2 [18], 問題 j3 [18]

ともに、正答の「11」、英文中に登場する「8」「14」を除くと「9」と「6」の選択率が高い。「9」は図中の数字17から8を引くと出てくる数字であり、「6」は14から8を引くと出てくる数字である。英文を読み取れないものの、何とか正しそうな数字を選択しようとしていた様子がうかがえる。

4 考察

4.1 設問ごとの分析結果について

本稿では、「言語運用力」試作問題として出題された六つの問題項目について、モニター調査で得られた結果を用いて項目分析を行い、選択枝の違いによる影響をみてきた。

問題 m では、選択枝を五つから四つに減らした結果、すべての誤答選択枝が5%以上選択されるようになった。正答率はほぼ同じなので、誤答選択枝が一つ減った分、残りの誤答選択枝が選ばれるようになったためと考えられる。多枝選択式形式では誤答選択枝もある程度選択されることが期待されている。したがって、今回のように選択枝を一つ除いても正答率や識別力などの項目の特性が変化しないのであれば、有効でない選択枝を省いた今回の変更は、選択枝の選択率の面からみても、受験生の読む分量を減らして負担を軽くするという面からみても、問題の改良につながる変更と言えよう。

問題 b では、選択枝の内容を工夫した結果、受験生がどこで間違えるのかを把握できるようになった。今回の変更によって、問題が何を測定しているのか、つまり正解した人と間違えた人とで何が理解できていて、何が理解できていないかの違いが分かるようになったという点で問題は改良されたと言える。正答できなかった場合でも、受験者本人の勉強になるだけでなく、たとえば学校の教師が授業中に指導する際にも役立つ情報が得られる。ただし、今回の変更では正答率も識別力も下がってしまったため、試験問題を構成する一つの問題としては改良されたとは言い難い。例えばヒントを加えたり、計算する必要性を無くしたりするなどして、想定する受験者層に合うように問題の困難度を下げる必要があるであろう。

問題 p では、選択率の低い選択枝を変更しても、

変更後も選択率が低い場合には、その他の選択枝へ影響しないことが分かった。これは逆に他の選択枝が適切であり、正解が分からない人にはそれぞれ魅力的に思える選択枝であったということの意味する。また、変更後の選択枝は、当初の選択枝よりは正答に見えやすいことを意図して変更したものであったが、それでも選択率は非常に低かったことから、正答に見えやすくはなかったと思われる。したがって次の変更としては、より魅力的な誤答選択枝にすることが考えられる。正答率を高くするための変更としては、椎名他(2014)にあるように想定する受験者層が取り組みやすくなるよう設問の意味を分かりやすく説明するなど考えられる。

問題 q では、選択枝を六つから四つに減らしても、正答率や識別力などに影響を与えない場合があることが示された。問題 q では選択枝が六つであっても四つであってもすべての誤答選択枝が5%以上選択されているので、どちらも問題としての性能はそれほど変わらないと言えよう。ただし、問題の分量が短くて済むのであれば十分な時間を確保するという意味で選択枝数を4にした方が良いと考えられる。

問題 j では、選択枝の内容が、内容によっては正解へのヒントになることが示された。問題 j2 [17] で正答率が高かったのは、選びやすい(英文中に登場する)数字が「8」しか選択枝中になかったためと思われる。しかし、全員が単純に「8」を選んだのではなく、上位の群ほど選択率が高く、問題の識別力は高い結果となっている。この問題では問題 j2の方が問題の正答率、識別力とも良い結果となっているが、一般的には、選択枝が正解へのヒントとなることは避けるべきであり、問題 j3へと変更する方が妥当であろう。また問題 j では、人数が少なくなることから高校生と大学生を合わせて分析を行った。試験を開発するに当たっては、想定する受験者層に近い集団(大学生のみ)を対象とした分析を今後行う必要がある。

本稿の分析により、問題を改良につながる変更や、今後の変更の指針が明らかとなった。今後は、椎名他(2014)、伊藤他(2014)での分析結果と併せて、さらに改良を加えていくこととなる^{注5)}。

4.2 問題の改良に向けて

前節でみたように、選択枝の変更によって、問題の特性（正答率や識別力など）は大きく変化する。また、問題 m, b, p, q で見たように、高校／大学のような受験者集団の違いによっても問題の特性は大きく変化する。問題を作成する際には、問題の正答率や識別力、選択枝の選択率などの問題の特性を適切に調整する必要があるが、これらの特性は受験者集団の能力によっても変化してしまう。そのため、本稿で行ったような正答率や識別力、設問解答率分析図など問題の特性に関する分析結果を蓄積したり、問題作成者の経験を重ねたりしていくことが、問題を改良していく際には必要となる。

しかし、テストの実施後に設問解答率分析図が公開されることは少なく、また、本稿のように選択枝の変更の前後で結果を比較したものが公開されることも少ない。本稿の分析は、「言語運用力」試作問題を検討するためのものであるが、ここで得られた知見の中には問題一般に通じるものがあるので、本稿の分析が様々な問題の作成にも役立てば幸いである。

注

- 1) これら6問の他に、本稿の分析対象とはしないが、選択枝ではなく設問中の図に変更を加えた大問が一つある。
- 2) 受験生が各誤答選択枝を選択する理由は、ここで挙げた理由だけには限らない。例えば、eachの意味が分からず、saveを「割り増しする」と誤って解釈した場合にも $5 \times 11 + 1 = 56$ ドルという解答に至る可能性がある。ここでは、saveとfreeの解釈がポイントになると考えた場合を論じている。
- 3) 無マーク、無効な選択枝番号へのマーク（問題q2の場合であれば、①～④以外の番号へのマーク）、多重マークを合わせて「無効な回答」とした。
- 4) 無効な回答の内訳は、無マークが7人、無効な選択枝番号へのマークが4人、多重マークが1人であった。
- 5) なお、改良された冊子を用いたモニター調査は2013年4月に実施された。この実施結果

については別の機会に報告する。

謝 辞

本稿を作成するにあたり有益なコメントをくださいました査読者に御礼申し上げます。また、モニター調査にご協力いただきました高校、大学、短期大学の教職員および各校の生徒、学生のみなさまに感謝いたします。

参考文献

- 大学入試センター (2003). 総合問題調査研究委員会調査研究報告書－国語と数学の基礎学力評価の試み.
- 大学入試センター (2004). 総合問題調査研究委員会調査研究報告書－総合基礎問題「国語」「数学」「英語」の作成－.
- 大学審議会 (2000). 大学入試の改善について (答申) http://www.mext.go.jp/b_menu/shingi/old_chukyo/old_daigaku_index/toushin/1315961.htm (2014年5月2日)
- Downing, S. M., & Haladyna, T. M. (2006). *Handbook of test development*. Lawrence Erlbaum Associates Publishers. (池田央監訳 (2008). 『テスト作成ハンドブック』株式会社教育測定研究所)
- Haladyna, T. M. (2004). *Developing and validating multiple-choice test items*. Lawrence Erlbaum Associates Publishers.
- 林篤裕 [研究代表者] 他 (2010). 平成18-20年度プロジェクト研究報告書新しい枠組みとしての総合試験の実証的研究, 大学入試センター研究開発部.
- Hogan, T. P. (2007). *Psychological testing: A practical introduction*. Hoboken, NJ: John Wiley & Sons. (繁榊算男・椎名久美子・石垣琢磨訳 (2010). 『心理テスト－理論と実践の架け橋－』培風館)
- 伊藤圭 [研究代表者] 他 (2011). 平成21-22年度プロジェクト研究I報告書新しい枠組みとしての総合試験の実証的研究 (2), 大学入試センター研究開発部.
- 伊藤圭・宮埜寿夫・椎名久美子・荒井清佳・桜井裕仁・田栗正章・小牧研一郎・安野史子 (2014). 大学入学志願者の基礎的学力測定のための英語問題の試作とモニター調査による予備的検討－正答率分析図を用いた問題内容と受験者の応答の事例分析－, 大学入試研究ジャーナル, **24**, 59-67.
- 桜井裕仁・田栗正章・安野史子・小牧研一郎・荒井清佳・伊藤圭・椎名久美子・宮埜寿夫 (2014). 大学入学志願者の基礎的学力測定のための「数理分析力」の調査とその予備的検討, 大学入試研究ジャーナル, **24**, 51-58.

- 椎名久美子・宮埜寿夫・伊藤圭・荒井清佳・桜井裕仁・小牧研一郎・田栗正章・安野史子 (2014). 大学入学志願者の基礎的学力測定のための枠組みの検討および「言語運用力」についての予備的分析, 大学入試研究ジャーナル, **24**, 41-49.
- 椎名久美子・杉澤武俊・櫻井捷海 (2007). 大学入試センター法科大学院適性試験の設計及び安定性に関する実証的検討, 日本テスト学会誌, **3**, 110-121.
- 杉澤武俊・内田照久・椎名久美子 (2009). 法科大学院適性試験が測定している能力・特性に関する実証的検討, 日本テスト学会誌, **5**, 128-135.
- 山村滋 (2010). 高校と大学の接続問題と今後の課題—高校教育の現状および大学に必要な技能の分析を通して—, 教育学研究, **77** (2), 157-170.
- 山村滋・荒牧草平 (2002). 高校教育多様化政策のもとでの受験シフト—普通科からの大学進学者の履修実態・受験実態の分析—, 大学入試研究ジャーナル, **12**, 101-107.
- 柳井晴夫 [研究代表者] 他 (2006). 平成 15-17 年度共同研究報告書総合試験問題の分析的研究, 大学入試センター研究開発部.

Analysis of data from trial practical reading test for measuring basic academic skills of university applicants. --- Effects of alternatives on item statistics ---

ARAI Sayaka*
ITO Kei*
SHIINA Kumiko*
MIYANO Hisao*
KOMAKI Ken-ichiro**
SAKURAI Hirohito*
TAGURI Masaaki***
YASUNO Fumiko****

Abstract

Prototypical items were produced for practical reading test for measuring basic academic skills of university applicants, and a trial for evaluating the items was planned and carried out. For this trial, four test forms having 6 common items with different sets of alternatives were developed. Statistical properties of these items were evaluated by using quintile item response charts. The results suggested that difference in alternatives affects item statistics to some extent depending on groups of examinees.

Key words: Practical reading, Multiple-choice, Item analysis,
Quintile item response chart

* Research Division, National Center for University Entrance Examinations
** National Center for University Entrance Examinations
*** Chuo University
**** National Institute for Educational Policy Research