

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 —

伊藤圭，宮埜寿夫，椎名久美子，荒井清佳，桜井裕仁（大学入試センター），
田栗正章（中央大学，大学入試センター），小牧研一郎（大学入試センター），
安野史子（国立教育政策研究所）

大学入学者選抜方法の多様化に対応するため，従来の教科・科目型試験とは異なる観点から受験者の能力を評価する「新しい試験」として，言語運用力と数理分析力の問題を試作し，問題評価のためのモニター調査を行った。本稿では，特に言語運用力の問題の一部として作成された英語問題について，問題内容と調査結果に基づいて分析した受験者の応答の特徴について報告する。分析の結果，試作問題は概ね調査参加者の能力に適した困難度であることが分かった。また，問題の選択枝を工夫することにより，問題内容を変更せずに困難度を操作できることが分かった。

1 はじめに

大学入学志願者の社会的，教育的背景の多様化に伴い，大学入学者選抜方法も多様化してきている。近年増加しているAO入試や推薦入試を経た大学入学者は，大学入試センター試験や個別大学の学科試験などの教科・科目型試験を受験していない場合が多く，これらの学生の学力を適切に評価する方法が求められている。

平成12年11月の大学審議会の最終答申「大学入試の改善について」では，幅広い層の大学進学に対応するため，教科・科目ごとの知識等ではなく，数理的な思考力や言語的な表現力，複数の教科・科目に基づく知識等を組み合わせて応用していく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た（大学審議会，

2000）。この間，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を行ってきた（柳井ほか，2006；林ほか，2010；伊藤ほか，2011）。また，平成15～22年度には，特定の専門分野の知識に依存せず，読解力，表現力，推論力，分析力等を評価するタイプの試験として，法科大学院適性試験が大学入試センターによって実施され，試験得点の統計的性質や測定している能力・特性の分析が行われている（椎名ほか，2007；杉澤ほか，2009）。

このような大学入学者選抜を取り巻く社会的要請および多面的，総合的な能力評価に関する先行研究の知見に基づき，大学入試センター研究開発部では，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平成23年度から進めている。その一環として，言語運用力を測定する問題と数理分析力を測定する問題を試作し，平成24年10月から同年12月までの

間に、高校 2 年生と大学・短大 1 年生を対象としたモニター調査を行った（椎名ほか、2013；桜井ほか、2013）。

本稿では言語運用力の問題の一部として出題した主に英語で提示された問題（以下、英語問題）に焦点を絞り、その問題内容とモニター調査のデータに基づく受験者の応答の事例分析を行った結果を報告する。

2 言語運用力問題の一部としての英語問題

今回試作した言語運用力の問題冊子は、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目指したもので、日本語で提示された問題（以下、日本語問題）と英語問題が 1 冊子あたりそれぞれ大問 4 問ずつ、合計 8 問の大問で構成されている。解答時間は 1 冊子 40 分である。

英語問題の作成にあたっては、一部の問題を除き、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、音声で提示されたリスニングの問題を文字による提示に置き換えたり、選択枝を一部変更したりした。大学入試センター試験の「英語リスニング」の問題は、「新しい試験」で想定する受験者層（AO 入試や推薦入試などを経て大学や短期大学に入学を志願する受験者層）にある程度適した難易度であると予想したためである。また、内容の面から見ても、大学入試センター試験の「英語リスニング」の問題は、日常に近い言語活動場面を題材にしており、音声を聞き取る能力以外にも実践的な能力との関係が深いと考えられる（荒井ほか、2009）ためである。表 1 に試作した問題の問題 ID（問題を識別するための記号）と試作に利用した問題の出典を示す。

一般に問題解決や課題遂行の場面では、新たな知識の獲得、既有知識との関連づけや新たな意味づけを行いながら思考し、その後の具体的な行動のための判断を行う。今回試作

表 1 試作英語問題の出典
（セ試：大学入試センター試験）

問題ID	出典
a	2009年度「セ試」英語リスニング第1問-問1
b1	2009年度「セ試」英語リスニング第1問-問6
b2	
c	2009年度「セ試」英語リスニング第2問-問7
d	2010年度「セ試」英語リスニング第2問-問8
e	2006年度「セ試」英語リスニング第3問-問14
f	2006年度「セ試」英語リスニング第4問-問21
g	2009年度「セ試」英語リスニング第1問-問3
h	2007年度「セ試」英語リスニング第2問-問12
i	2007年度「セ試」英語リスニング第2問-問10
j1	オリジナル問題
j2	
j3	
k	2006年度「セ試」英語リスニング第4問-問20

した英語問題がこのような実践的な能力を測定するものになっているかどうか、すなわち、問題の内容的妥当性を一通り確認するために、「単純想起」「意味の解釈」「判断」という 3 つの知的操作に対応づけが可能かどうかを検討した。これら 3 つの構成概念は、Bloom の教育目標分類（Bloom et al., 1956）のモデルを参考にして試験問題の分類をするために作成されたとされるもので、それぞれタキソノミー I 型、II 型、III 型などと呼ばれ（日本医学教育学会、2008）、医師国家試験や保健師助産師看護師国家試験等の試験問題を知識と実践的能力の観点から分類する際に用いられているものである。検討の結果、今回試作した英語問題についても、ある程度の対応づけが可能であることが分かった。

一方、日本語問題については、前節で述べた教科・科目フリー型総合試験や法科大学院適性試験等を参考にし、書かれていることを正確に読み取ったり、文章の文脈を理解したりする能力を測る問題の試作を試みた。参考

にした総合試験や適性試験は、学部レベルとしては困難度が高いが、いずれも特定の専門分野固有の知識ではなく、問題解決や課題遂行に必要な基本的な能力を測定するという点では本研究の試作問題と共通するものであり、問題の分量や論理構造などの複雑さの違いに注意すれば、本研究における問題の試作に十分応用可能である。

以上のような検討を経て、本研究においては、言語運用力とは一定の意味を有する言葉のまとまりであるテキスト（文章、または発話を文字として提示したもの）から必要な情報を読み取り、その情報を運用して知的活動に結びつける能力とした。日本語問題と英語問題のどちらも言語運用力を測定するという共通の枠組みの下で設計するために、上記の英語問題と対応づけた能力を包含しつつ、日本語問題とも統一的に対応づけることができる能力分類の枠組みを検討した結果、言語運用力の問題で測る能力を「L1: 情報の把握」「L2: 内容の理解」「L3: 推論と推察」という3つの構成概念にまとめた。

大学での学びはテキストを批判的に読んで自分の意見を構築したり、新たな問題を自ら発見したり、既存の知識だけでは解決できない問題に取り組んだりといった主体的なスタイルで行われる。これらの知的活動を行うにあたって、日本語や英語などの言語の違いに依らず、共通に必要な能力として定義し得る上記のL1～L3を言語運用力として測定することが本試作問題の目的である。

基本的な言語運用力を測るものであっても、日本語問題を解く上で日本語の知識は必要であり、その多寡がある程度成績に影響すると考えられる。同様に、英語問題を解く上で英語の知識が必要ないわけではない。しかし、本試作問題は個別の言語の固有知識の評価を意図するものではない。この点が教科科目の到達度を測る英語試験などと異なるところである。

3 問題冊子の種類と英語問題の配置

モニター調査で使用する試験問題冊子を4種類作成し、識別のために、それぞれにP, Q, R, Sという記号を割り当てた。これは、受験者属性や選択枝の変更が問題の解答に与える影響を評価できるようにすることと、できるだけ多くの試作問題の特性（困難度、識別力等）を把握できるようにするためである。問題冊子の種類と英語問題の配置を表2に示す。

表2に示されていない第1問、第4問、第5問、第7問は日本語問題である。aからkまでの記号は表1で定めた問題IDである。b1, b2やj1, j2, j3のように同じアルファベットと数字が組み合わさっている場合は、問題内容は同じであるが、与えられている図や選択枝の内容が異なることを表す。具体的な問題例は付録にまとめた。

表2 問題冊子中の英語問題の配置

問題番号	解答番号 PQR/S	P	Q	R	S	
第2問	問1	[3]/[5]	a	a	a	a
	問2	[4]/[6]	b1	b2	b2	g
第3問	問1	[5]/[7]	c	c	c	h
	問2	[6]/[8]	d	d	d	i
第6問	問1	[11]/[22]	e	e	e	e
	問2	[12]/[23]	f	f	f	f
第8問	[17][18]/[28]	j1	j2	j3	k	

4 問題冊子の割り当てと受験者数

モニター調査は、AO入試や推薦入試での大学進学者が比較的多い公立高校（H1）および一般入試での大学進学者が多い公立高校（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（いずれも私立；U1～U3）の1年生を対象として実施した。今回のモニター調査における問題冊子の割り当てと受験者数を表3に示す。

表 3 問題冊子の割り当てと受験者数

実施校	冊子の種類			
	P	Q	R	S
H1高校	71			
H2高校	76	39	36	79
H3高校	79	79	40	119
U1短大		34	33	36
U2大学		39	30	26
U3大学		38	35	40
合計	226	229	174	300

表 2, 3 から分かるとおり, 高校生のデータを用いることにより, P 冊子とその他の冊子の比較が可能である。また, 高校生と大学生の比較は主に Q, R 冊子を通して行うことができる。なお, 本稿では, 共通問題が多い P, Q, R 冊子のみを分析対象とし, S 冊子についての分析は省略する。また, H1 高校の調査は他の高校, 大学・短期大学の調査に先立って試行的に行われたものであり, 本稿での検討対象とはしない。

5 受験者の応答の事例分析

本節では, P, Q, R 冊子に含まれる問題に対する受験者の応答の事例分析の結果を述べる。以下に示す図 1~8 は項目応答の分析によく用いられる設問正答率分析図であり, 冊子の得点の順に受験者集団をほぼ等しい人数の 5 群に分け, 各受験者群の正答率を折れ線で表したものである。ここで L, LM, M, HM, H は, それぞれ下位群, 中下位群, 中位群, 中上位群, 上位群を表す。折れ線の縦軸方向の高低が困難度の高低を, 折れ線の傾きの大きさが識別力の大きさを意味する。

5.1 問題 a, c, d, e, f の分析

まず, P, Q, R 冊子に共通の英語問題

a, c, d, e, f の分析を行う。図 1~5 はそれぞれ問題 a, c, d, e, f の設問正答率分析図である。実線が高校生 (P, Q, R 冊子に解答した H2, H3 高校の生徒 349 名), 破線が大学生 (Q, R 冊子に解答した大学・短大生 209 名) を表す。これによって高校生と大学生の項目応答を比較することができる。

これら図 1~5 を見ると, 問題 a, c, d, e, f のいずれにおいても, 全体的に高

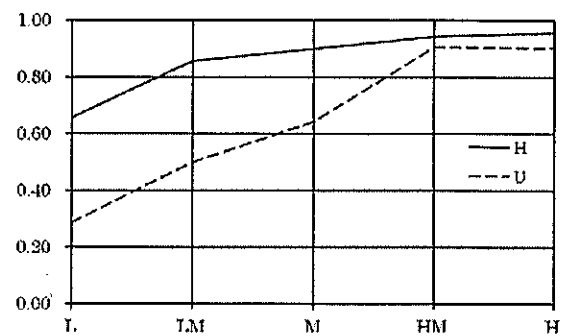


図 1 問題 a (第 2 問 - 問 1- [3])

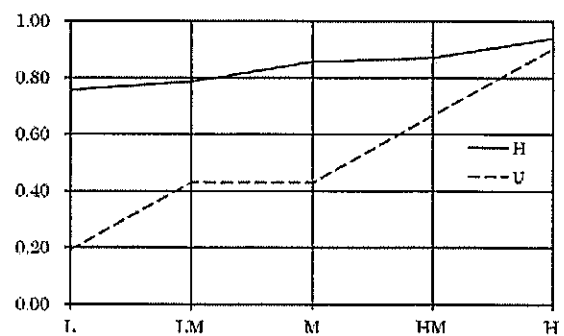


図 2 問題 c (第 3 問 - 問 1- [5])

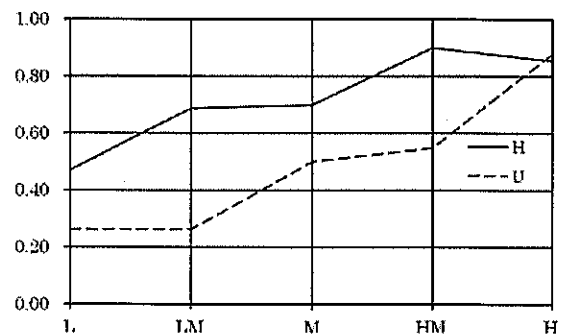


図 3 問題 d (第 3 問 - 問 2- [6])

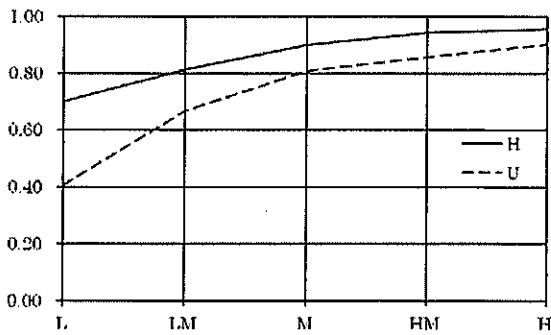


図4 問題 e (第6問 - 問1- [11])

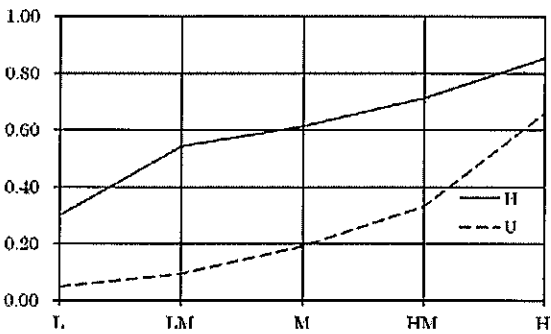


図5 問題 f (第6問 - 問2- [12])

校生の正答率の方が高く、高校生と大学生の正答率に開きがある。問題 a, c, d, e では高校生の識別力が大きいとは言えない。しかし、大学生では下位群に行くにしたがって正答率が下がり、上位と下位をよく識別している。

今回試作した問題は大学進学のために十分準備した受験生であれば、比較的高い正答率が得られるように目指したものであり、受験者全体(上位から下位まで)の正答率が高い場合には必然的に識別力が小さめの問題となり得るが、試験の目的に著しく反するものではない。実際に、高校生の参加者のうち半数程度は一般入試での大学進学者が多い高校からの参加者であった。問題 f では高校生と大学生のいずれにおいても比較的高い識別力が見られた。

基本的な能力を測る総合試験についての先行研究では、学科試験と総合試験の間に中程度の相関がみられ(伊藤, 2006), 教科科目

学力と基本的な能力とが互いに影響しながら両者の能力を向上させるものと思われる(伊藤ほか, 2010)。その意味では、今回の調査において、一般入試での大学進学者が多い高校からの参加者が含まれている高校生の正答率が大学生の正答率より高いことは不自然ではない。

既に大学入試を経た大学・短大生より、今後に大学入試を予定している高校生の方がテストに対する動機づけが強かったとも考えられるので、単純な比較はできないが、大学生の下位群での識別力が高かったのは高校での学習で十分に定着しなかった基礎的な能力が今回の試作問題で識別された結果とも考えられよう。

5.2 問題 b1, b2 の分析

次に P 冊子に収録されている英語問題 b1 と Q, R 冊子に収録されている英語問題 b2 の分析を行う。これによって同一内容の問題における選択枝の違いの効果を見ることが出来る。ここでは、P, Q, R 冊子に解答した H2, H3 高校の生徒 349 名のデータのみを用いる。問題 b1, b2 の具体的な内容は付録の A.1, A.2 節に示すとおりである。対話の内容から美術館の入場料を正しく計算させるものであり、save や free といった料金に関わる単語を正しく解釈できるかがポイントである。誤答選択枝はこれらのポイントに対応した起こり得る誤りのパターンからなる。

b1 より b2 の方で、より多くの誤りのパターンに対応した選択枝が用意されている。

受験者の解答結果から誤答選択枝の選択状況を確認したところ、b1 では、3つの誤答選択枝のうち2つが比較的选择率が高く、残りの1つは低かった。b2 では、b1 で选择率の低かった誤答選択枝は元々含まれておらず、新たに3つの誤答選択枝が追加されているが、追加した3つのうち1つだけが比較的高い选择率を有していた。したがって、b2

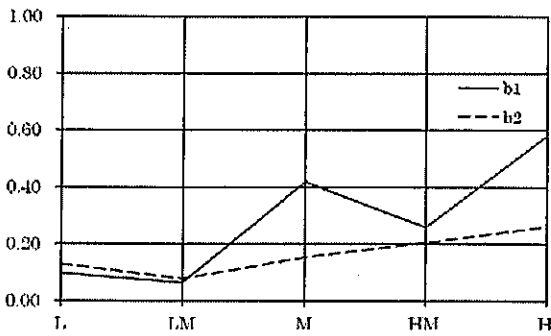


図 6 問題 b1, b2 (第 2 問 - 問 2 - [4])

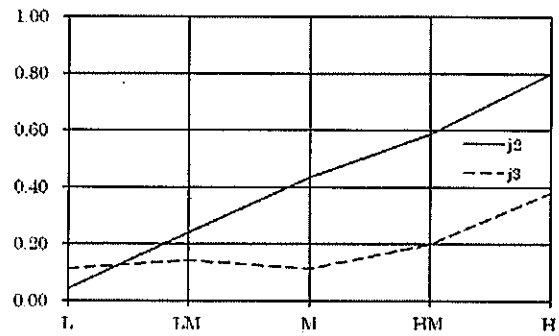


図 7 問題 j2, j3 (第 8 問 - [17])

では元々 b1 にあった 2 つの魅力的な誤答選択枝にさらに 1 つの魅力的な誤答選択枝が加えられたことになる。図 6 を見ると b2 (破線) の正答率の方が b1 (実線) の正答率より低く、魅力的な誤答選択枝の追加が困難度を高めていると推察される。また、対話内容の理解に加え、それに基づいて計算をするという複雑な操作が必要になるため、b1, b2 のどちらも全体として比較的困難度が高くなったものと思われる。

5.3 問題 j2, j3 の分析

最後に Q 冊子に収録されている英語問題 j2 と R 冊子に収録されている英語問題 j3 の分析を行う。これも同一内容の問題における選択枝の違いの効果を見ることができる。ここでは、Q, R 冊子に解答した H2, H3 高校の生徒 194 名、大学・短大生 209 名、合計 403 名のデータを用いる。問題 j2, j3 の具体的な内容は付録の A.3, A.4 節に示すとおりである。j2 と j3 の違いは選択枝の内容である。

j2 の誤答選択枝は連続する番号を機械的に並べただけであるが、j3 の誤答選択枝は起こりそうな誤りを意図的に配置したものである。そのため、j2 に比べて j3 の方が受験者を惑わす効果が大きく、図 7 に示すように j3 の方が正答率が低くなったと思われる。

次に、図 8 を見ると、最上位群 (H) だけ j3 より j2 の正答率が高いことが顕著である。解答番号 [17] では正答となる数値「8」

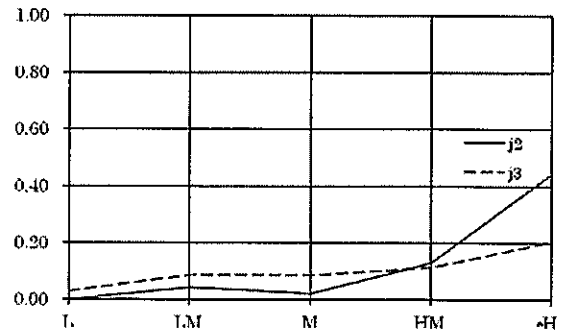


図 8 問題 j2, j3 (第 8 問 - [18])

が問題文中に示されており、問題文を正しく読み取れば正答することができるのに対して、解答番号 [18] では図中で伏せられている数値を計算で求め、さらに樹状図の分岐の一段前の数値から減算をするという過程が必要となるため、最上位群以外では困難度が高すぎて j2 と j3 で差が生じなかったものと思われる。

さらに誤答選択枝の選択状況を確認したところ、j2, j3 とともに「選択枝番号」と「図中の空欄に入る数」を混同していると思われる答案が見られた。また、j3 では、j2 に存在しない選択枝「⑥ 14」が比較的多く選択されていた。これは問題文中に「14」という数値があるため、問題文の読解が不十分な受験者が誤って選択したものと思われる。

6 おわりに

今回、900 名を超える受験者に対して、4 種類の問題冊子を割り当てることにより、多

くの試作問題の特性を把握することができた。その結果、試作問題は概ね「新しい試験」が想定する受験者の能力レベルに合った困難度であったと思われる。

また、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することができることが分かった。

今後、選択枝の影響をより詳細に分析するには、成績順の受験者群ごとに選択枝の選択率を見ていく必要がある。また、試験の妥当性をより詳しく検証するには、試作した「新しい試験」と基本的な能力・資質に関する外的基準との関連を調べるが必要と思われる。

なお、平成 25 年 4 月に、本稿で述べたモニター調査の結果に基づいて改良された試験問題を用いて、新たに、より規模の大きなモニター調査を行った。その結果については別途報告する予定である。

参考文献

- 荒井清佳・橋本貴充・杉澤武俊・荘島宏二郎・伊藤圭 (2009). 「センター試験『英語』得点と受験者の主観評価に基づく英語能力イメージとの比較」『日本テスト学会第 7 回大会発表論文抄録集』 114-115.
- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., and Krathwohl, D. R. (Eds.) (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain*. New York: David McKay.
- 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 >

(2013 年 9 月 20 日)

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.

伊藤圭 (2006). 「総合試験と教科・科目別試験との関連についての分析」『平成 18 年度大学入試センター試験モニター調査研究報告書』大学入試センター研究開発部, 89-103.

伊藤圭・林篤裕・椎名久美子・田栗正章・小牧研一郎・柳井晴夫 (2010). 「学科試験および科目得意度との比較による総合試験の妥当性の検証」『日本テスト学会誌』 6, 113-124.

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.

日本医学教育学会 (監修) (2008). 『医学教育マニュアル 4 「評価と試験」』篠原出版社.

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・杉澤武俊・桜井捷海 (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.

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 .
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.
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 .

A 付録 (問題例)

A.1 問題 b1 (第 2 問 - 問 2)

問 2 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
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

A.2 問題 b2 (第 2 問 - 問 2)

問 2 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
入場料は何ドルか。正しいものを、下の①～⑥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 44 ドル ② 52 ドル ③ 55 ドル ④ 56 ドル ⑤ 65 ドル ⑥ 70 ドル

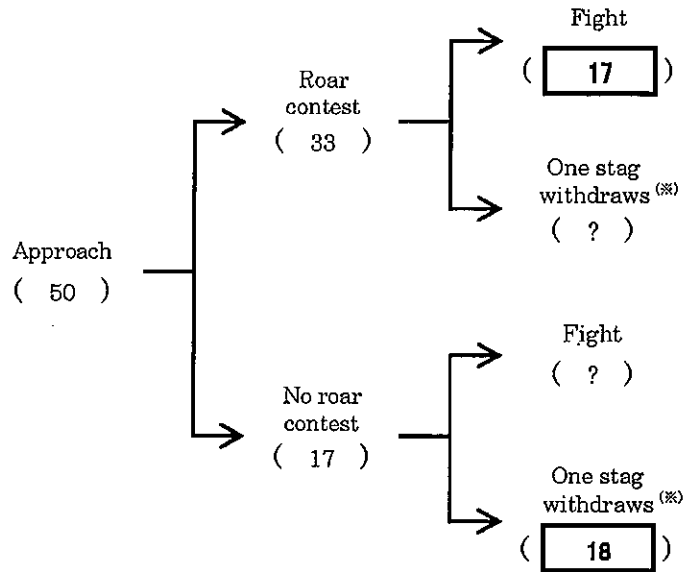
A.3 問題 j2 (第 8 問)

雄鹿 (stag) は、9 月の終わり頃から 10 月の初め頃にかけて、雄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 ・ 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。

(次頁に続く)

(前頁からの続き)

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.



※ withdraw (引き下がる)

Figure 1

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

A.4 問題J3 (第8問)

雄鹿 (stag) は、9月の終わり頃から10月の初め頃にかけて、雄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 ・ 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.

(A.3と同じ図)

Figure 1

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