

平成 30 年 1 月 9 日

独立行政法人 大学入試センター  
理事長 山本 廣基 殿

日本テスト学会  
理事長 繁榊算男

平成 29 年 11 月実施の大学入学共通テスト導入に向けた試行調査に関する  
日本テスト学会の意見

貴センターは、平成 29 年 12 月 4 日付で、平成 29 年 11 月 13 日～24 日に実施された、大学入学共通テストの導入に向けた試行調査（プレテスト）の概要、問題及び正答率の速報等を公表されました。

平成 32 年度から実施が予定されている大学入学共通テストがよりよきものとなるよう、以下に、試行調査で公開されたテスト問題の改善の方策等について日本テスト学会としての意見を述べます。

本来であれば試行調査のデータの詳細な分析に基づいて意見を申し述べるべきところですが、データそのものは公開されていないため、テスト問題の形式的側面や、貴センターから公表された正答率等、限られた資料のみからの意見となることをあらかじめお断りします。

また、以下に述べることは、貴センターにおいては十分に承知されている内容も多いことと思いますが、本意見を日本テスト学会のホームページ等で公表して、一般の方にも大学入学共通テストの今後のあり方について広く理解していただくことも目的としていますので、初歩的あるいは冗長と思われる箇所につきましても、趣旨をご理解いただき、ご寛恕いただければと存じます。

意見 1：テストの分量と設問数について

大学入試における共通テストの受験者は幅広い学力層からなるため、十分な識別力を確保するためには、難易度の高い問題から易しい問題まで、万遍なく出題されることが求められます。また、出題範囲が偏ることによって、たまたまその部分だけ学習してきた受験者が有利になる状況、いわゆる「ヤマが当たる」状況を避け、内容的妥当性と信頼性を確保するためにも、指導要領及び教科書を基礎にした出題範囲から広く出題されることが求められます。そのためには、十分な数の設問が必要となります。

一方、試験時間は限られていますので、試験時間に対して設問数が多すぎると、受験者の反応の速さを測定することを目的とした「スピードテスト」となってしまう、本来測るべき学力とは異なるものを測ることになりかねません。テストは一定の制限時間が設けられる

一方で、受験者の力量を十分に測ることが要請されますから、多数の受験者が時間切れのためすべての問題に解答できないような出題量は望ましくありません。

テストはこの両方の要請に応えるべくバランスをとる必要がありますが、もし設問数の割に読ませる文章や資料が過度に多くなると、試験時間の多くがこれらの文章や資料を読むことに使われるため、一定の時間内に問うことのできる設問が少なくなってしまいます。

今回の試行調査の問題は、添付の【補足資料】に示すように、多くの科目において設問あたりの読ませる文章・資料の量が多くなっており、その分、設問数が減少しているため、テストの識別力、内容的妥当性、信頼性の低下が懸念されます。提示する文章・資料が、それぞれの科目で測りたい学力を測るうえで必須の内容かどうか精査し、厳選することが求められます。

### 意見2：選択解答の形式について

多肢選択問題は「択一式」、すなわち適当な選択肢を1つ選ぶのが基本ですが、今回の試行調査では、「適当なものをすべて選べ」という解答形式が目立ちます。この解答形式は、「適当なものがいくつあるのか」もあわせて問うことになるため、より深い思考を求めていると誤解されることがありますが、実際には、1つ1つの選択肢について「適当か否かの二者択一」をしているにすぎず、特に異なる力を測定するものではありません。

「適当なものをすべて選べ」という解答形式によって生じる変化は、正答率の低下です。これは添付の【補足資料】に示されている通りです。関連して生じるもう1つの変化は、識別力の低下です。たとえば「5つのうちから適当なものをすべて選べ」という設問は、上述のように5つの二者択一問題と同等であり、二者択一にすれば5問正答から0問正答までの細かい個人差を見ることができものを、5問正答のみを正答し、4問以下の正答は0問正答と同じとみなすわけですから、貴重な個人差情報を捨ててしまうこととなります。

また、テスト理論上のこととは別に、マークシート解答の読み取りにおいても、複数の選択肢を選ばせることはトラブルのもとになることが懸念されます。それは受験者がマークを消して付け直すことをした場合、消し方によっては、消したのかどうか判然としないことがあり、その判断が、1つのみマークすることになっている場合に比べ、いくつでもマークしてよいことになっている場合にはより難しくなると考えられるからです。

このように、「適当なものをすべて選べ」という解答形式は、単に正答率を低下させ、識別力を低下させる機能しかなく、採点実務上のトラブルにつながる可能性もあるため、原則として使用しないことが望まれます。

### 意見3：記述式問題について

記述式問題は選択式問題に比べ、解答にかなりの時間を要するため、意見1で述べた試験時間の観点からすると、導入するならば記述式問題ならではの質の高い問題でなければならぬということが言えます。ただし、ここでは今回の試行調査の問題がその要請に応える内

容のものであるかどうかの判断は控え、採点の信頼性とフィージビリティの観点からのみ意見を述べます。

試行調査の記述式問題は、採点の信頼性を重視して、また自己採点も可能となるよう、条件付き記述式の形式をとっており、採点基準が細かく提示されています。ただ、この形式と採点基準で、どれだけ正確な採点が可能か、つまり採点者間で一致し、自己採点とも一致する採点となるかは採点データの分析を待って判断することになります。

また、フィージビリティについては、採点にかかる時間の問題があります。そして、この問題は上述の採点の信頼性の問題とも関係します。十分な時間があれば、採点者間で判断のずれが生じた場合の調整も十分にできるはずですが、時間が切迫している場合、そのこと自体が採点作業に影響し、採点の信頼性を損ねることにつながる心配があります。

したがって、現在想定されている試験実施から結果報告までの期間内で、余裕をもって正確な採点が可能であることを試行調査のデータから示すことが求められます。

以上、主な3点のみに絞って意見を申し上げました。今後の参考にしていただければ幸いです。

以上

### 【補足資料】

ここでは、意見 1 および意見 2 を補足する資料として、今回の試行調査および過去 3 年間の大学入試センター試験（以下、センター試験）の設問数、ページ数、1 ページ当たりの設問数（設問数/頁）、平均正答率を示しています。ただし、試行調査が実施された科目のうち、記述式が加わって大きく条件が変わった「国語」及び「数学Ⅰ・数学A」については除いています。

「設問数/頁」は試験問題のコンパクトさのおおまかな指標として算出したものです（複数のマークの組み合わせで構成されている設問があるため、マーク数が設問数を上回っているケースがあります）。また、「平均正答率」は、試行調査は各項目の得点率を平均した「平均得点率」、センター試験はセンター公表の「平均点」を用いています。

平均正答率の比較については、試行調査が本番の入試とは異なる状況で行われること、実施の時期および対象者の学年も同じでないこと、また、試行調査の対象者層がセンター試験の受験者層をどの程度近似できるかについての資料がないことから、参考程度の資料となることをあらかじめお断りしておきます。

#### (1) 数学Ⅱ・数学B

表 1 に「数学Ⅱ・数学B」についての値を示します。「数学Ⅱ・数学B」の場合、ページ数が約 1.6 倍に増えているにもかかわらず、設問数が約 3 割減っています。結果として、「設問数/頁」は 1/2 以下になっています。平均正答率は 40% 強です。

表 1. 数学Ⅱ・数学B

	設問数	ページ数	設問数/頁	平均正答率（得点率）
試行調査	36.3	18	2.02	41.9%
H27 センター試験	50.0	11.3	4.41	39.3%
H28 センター試験	55.7	11.3	4.91	47.9%
H29 センター試験	46.7	11.3	4.12	52.1%
センター試験平均	50.8	11.3	4.48	46.4%

\*3 問中 2 問を選択解答する選択問題があるため、ページ数、設問数はそれぞれ 3 つの選択パターンの平均値を取っています

#### (2) 地理歴史・公民

表 2 に「地理歴史・公民」についての値を示します。「世界史B」「日本史B」「地理B」に共通の特徴は、問題冊子のページ数が大幅に増えた（約 1.2～1.7 倍）ことです。設問数はほぼ同じですので、受験者にとっては問題文や資料の読み込みの負担が増えたこととなります。平均正答率も低めで、「スピードテスト」化が懸念されます。

一方、「現代社会」はページ数を増やすことなく、その分を設問数の減少で対応しています。時間の制約についての1つの可能な解決策にも見えますが、設問数が減るとカバーできる領域や学力層が薄くなり、識別力・内容的妥当性・信頼性の低下が懸念されます。また、「地理歴史・公民」の中での科目間で方針が異なるのも望ましいことではありません。

方針は異なるものの、いずれの科目も結果として「設問数/頁」が3~4割も少なくなっており、試験としてのコンパクトさ、測定効率の低下も懸念されます。

表2. 地理歴史・公民（世界史B、日本史B、地理B、現代社会）

		設問数	ページ数	設問数/頁	平均正答率（得点率）
世界史B	試行調査	36	45	0.80	53.0%
	H27 センター試験	36	24	1.50	65.6%
	H28 センター試験	36	28	1.29	67.3%
	H29 センター試験	36	27	1.33	65.4%
	センター試験平均	36.0	26.3	1.37	66.1%
日本史B	試行調査	30	36	0.83	48.8%
	H27 センター試験	36	30	1.20	62.0%
	H28 センター試験	36	30	1.20	65.6%
	H29 センター試験	36	30	1.20	59.3%
	センター試験平均	36.0	30.0	1.20	62.3%
地理B	試行調査	30	43	0.70	51.3%
	H27 センター試験	36	34	1.06	58.6%
	H28 センター試験	35	34	1.03	60.1%
	H29 センター試験	35	34	1.03	62.3%
	センター試験平均	35.3	34.0	1.04	60.3%
現代社会	試行調査	23	34	0.68	51.2%
	H27 センター試験	36	36	1.00	59.0%
	H28 センター試験	36	32	1.13	54.5%
	H29 センター試験	36	34	1.06	57.4%
	センター試験平均	36.0	34.0	1.06	57.0%

### (3) 理科

表3に「理科」についての値を示します。「物理」は設問数もページ数もほぼ変化なしですが、「化学」ではページ数は変わらずに設問数が減少しています。一方、「生物」と「地学」はページ数が増えて設問数が減少しており、ここでも科目によって傾向が分かれています。

「物理」「化学」「生物」に共通の特徴は平均正答率の低さです。受験者層が本番とは異なる

るにしても、センター試験と比較して20～30ポイントも下がっているのは、難問化の傾向があると言えるでしょう。

試験問題の内容そのものに原因がある可能性もありますが、顕著なのが意見2で述べた解答形式の特徴です。特に「物理」で、正解の選択肢数を示さずに選ばせ、すべて正しい場合のみを正答とする設問が多く、それらの設問はおおむね正答率が低くなっています。

地学では「すべて正しい場合のみを正答」とする設問が1問含まれており、その正答率は低いものの、1問のみであったため、科目の平均正答率には大きな影響が見られません。

表3. 理科（物理、化学、生物、地学）

		設問数	ページ数	設問数/頁	平均正答率（得点率）
物理	試行調査	21	23	0.91	31.9%
	H27 センター試験	21.0	21.5	0.98	61.6%
	H28 センター試験	21.0	26.5	0.79	61.7%
	H29 センター試験	23.0	29.5	0.78	62.9%
	センター試験平均	21.7	25.8	0.85	62.1%
化学	試行調査	25	27	0.93	35.0%
	H27 センター試験	29.0	26.5	1.09	59.2%
	H28 センター試験	29.5	25.0	1.18	54.5%
	H29 センター試験	35.0	27.0	1.30	51.9%
	センター試験平均	31.2	26.2	1.19	55.2%
生物	試行調査	28	36	0.78	34.6%
	H27 センター試験	31.5	31.5	1.00	48.4%
	H28 センター試験	32.0	31.0	1.03	63.6%
	H29 センター試験	34.5	34.0	1.01	69.0%
	センター試験平均	32.7	32.2	1.02	60.3%
地学	試行調査	28	35	0.80	44.1%
	H27 センター試験	30.0	29.5	1.02	40.9%
	H28 センター試験	30.0	30.0	1.00	38.6%
	H29 センター試験	30.0	23.0	1.30	53.8%
	センター試験平均	30.0	27.5	1.11	44.4%

\*センター試験においては選択問題があるため、ページ数、設問数はそれぞれ2つの選択パターンの平均値を取っています。

\*\*平成27年度センター試験「物理」「化学」「生物」は受験者数530,126名（99.98%）を対象とした平成27年1月23日発表（中間集計その2）に基づく「調整前」の数値です。

【主な参考文献】

- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Brennan, R. L. (Ed.) (2006). *Educational measurement* (4th ed.). American Council on Education.
- Lane, S., Raymond, M. R., and Haladyna, T. M. (Eds.). (2015). *Handbook of test development* (2nd ed.). Routledge.
- 日本テスト学会 (編) (2007). テスト・スタンダード—日本のテストの将来に向けて 金子書房
- 日本テスト学会 (編) (2010). 見直そう, テストを支える基本の技術と教育 金子書房