

平成 27 年度
大学入学希望者学力評価テスト（仮称）
フイージビリティ検証事業 総合報告書

平成 28 年 3 月
独立行政法人 大学入試センター 研究開発部

研究組織

a. 大学入試過去問の分析及びその新テスト問題への活用

研究代表者: 荒井 清佳 大学入試センター研究開発部 助教
研究分担者: 石岡 恒憲 大学入試センター研究開発部 教授
研究協力者: 平成24年度試験問題作成部会 部会長及び副部会長(非公開)
平成25年度試験問題作成部会 部会長及び副部会長(非公開)
平成26年度試験問題作成部会 部会長及び副部会長(非公開)

b. 新テスト試作問題の作成と評価

全体統括: 大津 起夫 大学入試センター研究開発部 教授(研究開発部長)

b-1 作業グループによる新テスト試作問題の作成

b-2 試作問題評価のためのモニター調査の実施

b-1統括: 椎名 久美子 大学入試センター研究開発部 教授
b-2統括: 内田 照久 大学入試センター研究開発部 准教授

大学入学希望者学力評価テスト(仮称)作業チーム(非公開)

試作問題の作成のための作業部会オブザーバー(国語, 数学, 地理歴史(世界史), 物理)

椎名 久美子 大学入試センター研究開発部 教授
桜井 裕仁 大学入試センター研究開発部 准教授
荒井 清佳 大学入試センター研究開発部 助教
伊藤 圭 大学入試センター研究開発部 准教授

英語の試作問題の要件のまとめ

立脇 洋介 大学入試センター研究開発部 助教

モニター調査の立案・渉外・実施

内田 照久 大学入試センター研究開発部 准教授
立脇 洋介 大学入試センター研究開発部 助教
桜井 裕仁 大学入試センター研究開発部 准教授
橋本 貴充 大学入試センター研究開発部 助教
荘島 宏二郎 大学入試センター研究開発部 准教授
石岡 恒憲 大学入試センター研究開発部 教授

モニター調査の分析

椎名 久美子 大学入試センター研究開発部 教授
桜井 裕仁 大学入試センター研究開発部 准教授
荒井 清佳 大学入試センター研究開発部 助教
伊藤 圭 大学入試センター研究開発部 准教授
立脇 洋介 大学入試センター研究開発部 助教
橋本 貴充 大学入試センター研究開発部 助教

事務担当: 事業部事業第一課企画調査係

里見 康弘
青木 智美
大野 純一
山田 華子(事務補佐員)
小倉 操子(事務補佐員)

モニター調査の準備補助

重高 史(事務補佐員)
大村 恵以子(事務補佐員)
小野塚 暁子(事務補佐員)

c. 新テスト実施環境の開発

c-1 CBT実施のプロトタイプシステムの開発

c-2 タブレットコンピュータを用いた障がいのある者のための受験上の配慮方法の検討

c-1担当: 大久保 智哉 大学入試センター研究開発部 助教
c-2担当: 南谷 和範 大学入試センター研究開発部 准教授
立脇 洋介 大学入試センター研究開発部 助教

平成27年度 大学入学希望者学力評価テスト(仮称)フィージビリティ検証事業
総合報告書

目次

研究組織	i
I 研究の目的	1
	大津起夫
II 各課題の研究成果のまとめ	
a 大学入試過去問の分析及びその新テスト問題への活用	3
	荒井清佳・石岡恒憲
b 新テスト試作問題の作成と評価	
1. 「b. 新テスト試作問題の作成と評価」の概要	17
	大津起夫・椎名久美子・内田照久
2. 各作業チームによる問題イメージの検討	19
	椎名久美子・桜井裕仁・荒井清佳・伊藤圭・大津起夫
3. 英語4技能を測定する問題イメージの作成	27
	立脇洋介
4. 試作問題評価のためのモニター調査の実施	37
	内田照久・立脇洋介・桜井裕仁・橋本貴充・荘島宏二郎・石岡恒憲
5. モニター調査の分析結果	
5.1 国語	45
	椎名久美子・橋本貴充
5.2 数学	52
	桜井裕仁
5.3 世界史	60
	荒井清佳
5.4 物理	66
	作業チーム(理科部会)・伊藤圭
5.5 試作問題の特徴を理解するための研究	73
	立脇洋介
c 新テスト実施環境の開発	
c-1 CBT実施のプロトタイプシステムの開発	87
	大久保智哉
c-2 タブレットコンピュータを用いた障がいのある者のための 受験上の配慮方法の検討	113
	南谷和範・立脇洋介

I 研究の目的

1. 研究の目的

本事業は、中教審答申「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」（平成 26 年 12 月 22 日）において示された「大学入学希望者学力評価テスト（仮称）」の実現に向けて、適切な試験問題の内容と実施方法についての知見を得るためモデル問題を作成し、それらの信頼性・妥当性と実現可能性について検討を加えるとともに、これら実施における CBT の利用可能性について検討することを目的として実施されたものである。

上記の中教審答申の提案内容は、その後、文部科学省が示した「高大接続改革実行プラン」（平成 27 年 1 月 16 日）によって設置が定められた専門家会議「高大接続システム改革会議」において引き続き検討されることとなった。本事業においては、「高大接続システム改革会議」での議論の進展を考慮しつつ、来るべき時代にふさわしい試験問題の作成を試みるとともに、それらの実施に係わる実証的検討を行った。

研究は、a. 「大学入試過去問の分析およびその新テスト問題への活用」、b. 「新テスト試作問題の作成と評価」、c. 「新テスト実施環境の開発」の 3 つのテーマとそれらのサブテーマから構成されるが、いずれの研究においても、大学入学者選抜をこれからの時代にふさわしいものに適合させるための試みの成果が示されている。新たな大学入学試験の実現のためには様々な進展が必要となるが、本研究がその第一歩となることができるならば幸いである。

本報告書のまとめには、主として大学入試センター研究開発部の教員があたったが、事業の実施はより多くの方々の協力によって可能になった。作業グループに参加していただいた各教科科目分野の諸先生方をはじめ、文部科学省の関係各位、大学入試センター総務企画部および事業部関係スタッフの尽力によって、本報告の完成に辿り着くことができた。関係各位に深く感謝する。

II 各課題の研究成果のまとめ

a 大学入試過去問の分析及びその新テスト問題への活用

1 研究実施状況

検証事業 a では、これまでに実施された大学入試問題等が測定している内容について思考力・判断力・表現力等がどのように問われているかの分析を行った。具体的には、大学入試センター試験（以下、センター試験）の主要 5 教科および法科大学院適性試験等の問題を対象とした。対象とした科目は下に示す。分析によっては、このうちのいくつかの科目のみを対象としている。

問題ごとに分析結果を整理し、統計情報を加えた上でまとめることで、新テストの問題作成のための支援素材として活用化することを目的とする。

大学入試センター試験

5 教科・11 科目（平成 24～26 年度、いずれも本試験）

- 英語：英語筆記
- 数学：数学 I・数学 A，数学 II・数学 B，情報関係基礎
- 国語：国語（第 1 問及び第 2 問）
- 理科：物理 I，化学 I，生物 I
- 社会：世界史 B，日本史 B，地理 B

法科大学院適性試験

2 科目（平成 20～22 年度、いずれも本試験）

- 第 1 部（推論・分析力）
- 第 2 部（読解・表現力）

本検証事業は、1) 外部評価者による評価，2) 試験問題評価委員会報告書からの抜粋，3) 作題部会部会長らによる試験問題総括，4) センター試験の主要科目のまとめ，の四つから構成される。これらの実施状況は次の通りである。

- 「1) 外部評価者による評価」では、試験問題について詳しい専門家にセンター試験や法科大学院適性試験の各設問がどのような問題であるかの評価を依頼し、各科目 2 名の外部評価者からの回答を得た。
- 「2) 試験問題評価委員会報告書からの抜粋」では、センター試験の計 11 科目，3 年度分の本試験の各設問についての評価を「試験問題評価委員会報告書」から抜粋した。それらと各設問の正答率や項目相関（統計情報）をまとめた。
- 「3) 作題部会部会長らによる試験問題総括」では、1) 及び 2) の資料を提示した上でセンター試験の問題作題部会の部会長らに担当した試験問題の総括を依頼し、その回答をまとめた。
- 「4) センター試験の主要科目のまとめ」では、センター試験の主要科目に焦点を当て、1) ～3) の評価結果を元に、センター試験で測っている内容の分析を行った。内容の分析だけでなく、歴史的な変遷や形式面の分析も行った。

2 研究成果の要約

本検証事業では、1) 外部評価者による評価、2) 試験問題評価委員会報告書からの抜粋、3) 作題部会部長らによる試験問題総括、4) センター試験の主要科目のまとめを行った。

2) の「試験問題評価委員会報告書」には、高等学校教科担当教員の意見・評価、教育研究団体の意見・評価、問題作成部会の見解がまとめられている。それらの評価に加えて、1) において当該科目の試験問題に詳しい専門家による外部評価を行い、さらに3) において作題部会部長らによる試験問題の総括を行い、多方面からの評価を得ることができた。4) は、これらの評価結果を元にセンター試験の分析を行ったものである。

また、本検証事業の成果については、必要に応じて検証事業 b-1 の作業チームに情報提供を行った。

以下では、各評価で得られた成果の要約を示す。

2.1 「1) 外部評価者による評価」の成果

外部評価者による評価では、各科目 2 名の外部評価者からの回答を得た。計 13 科目、3 年度分の本試験のすべての問題について網羅的な評価を得た。これにより、知識を適用したり、情報を抽出したり、趣旨を把握したりするなどの力を問うような問題も出題されていることが示された。また、同じ科目内の大問であっても、大問ごとに測ろうとする能力に違いがあることが示された。

ここでは、主要科目である「国語」「数学 I・数学 A」「数学 II・数学 B」の 3 科目について、外部評価者による評価のまとめを行った結果を示す。

外部評価者による評価観点のうち、全 13 科目に共通するものは「『論理的な思考』6 ラベル」¹ と「オーストラリアの 26 スキル」² の二つである。そこで、この二つの評価観点について、外部評価者が各設問に付けたラベルのうち、2 名の評価が一致したものを抜き出し、各科目、大問ごとに、各ラベルの付いた設問数及び配点の合計を整理して表にまとめた。ここに示す表は平成 26 年度試験の結果のみであるが、平成 24 年度試験、平成 25 年度試験も多少の違いはあるが同様の結果であった。

国語

国語についてまとめた結果を表 1 に示した。

「論理的な思考」6 ラベルでは第 1 問のみに「③把握・評価」「④関係・洞察」「⑥構造・判断」のラベルがついた。国語の第 1 問は「評論」、第 2 問は「小説」であるので、第 1 問にのみ「論理的な思考」に関係するラベルがついたのは妥当な結果である。

オーストラリアの 26 スキルについては、第 1 問、第 2 問ともに「2. 解釈」のラベルが多かった。ともに文章を読解する必要のある設問が多いことから、妥当な結果である。また、第 1 問に

¹ 「『論理的な思考』6 ラベル」とは、国立教育政策研究所が行った「特定の課題に関する調査（論理的な思考）調査」における六つの活動であり、① [理解・適用] 規則、定義、条件等を理解し適用する、② [抽出・分析] 必要な情報を抽出し、分析する、③ [把握・評価] 趣旨や主張を把握し、評価する、④ [関係・洞察] 事象の関係性について洞察する、⑤ [仮説・検証] 仮説を立て、検証する、⑥ [構造・判断] 議論や論証の構造を判断する、である（国立教育政策研究所、2013）。

² 「オーストラリアの 26 スキル」とは、オーストラリア、クイーンズランド州の高校生が身につけることが期待されているスキルのうち、山村（2010）によって 26 個に絞り込まれたものである。「1. 記憶」「2. 解釈」「3. 読む」などがある。

表 1: 外部評価者による評価のまとめ (平成 26 年度国語)

	「論理的な思考」6 ラベル			オーストラリアの 26 スキル※		
	ラベル	設問数	配点	ラベル	設問数	配点
第 1 問 (11 問 ・ 50 点)	③把握・評価	1 問	8 点	2. 解釈	2 問	12 点
	④関係・洞察	1 問	8 点	13. 関係	1 問	8 点
	⑥構造・判断	1 問	4 点	19. 分解	1 問	4 点
	なし	8 問	30 点	2. 解釈,13. 関係	1 問	8 点
				2. 解釈,13. 関係,24. 整理	1 問	8 点
			なし	5 問	10 点	
第 2 問 (9 問 ・ 50 点)	なし	9 問	50 点	2. 解釈	2 問	10 点
				2. 解釈,11. 感情	2 問	15 点
				2. 解釈,19. 分解	1 問	8 点
				2. 解釈,21. 判断	1 問	8 点
				なし	3 問	9 点

※オーストラリアの 26 スキルの内容

2. 解釈：言葉や他の記号（イラストなども含む）の意味を解釈する力

11. 感情：文章や人の考え方、絵画などに感情移入できること

13. 関係：アイデア・テーマ・問題などを相互に関係づけること

19. 分解：分析すること。ある物事を分解して、それを成立させている成分・要素・側面を明らかにすること

21. 判断：直面する状況に対して適切に判断・評価することができる力

24. 整理：必要な情報を探し出し整理する力

は「13. 関係」「19. 分解」「24. 整理」のラベルがつき、文章のテーマを関係付けたり、分析したり、必要な情報を探し出したりする力が求められていることが示された。一方、第 2 問には「11. 感情」「19. 分解」「21. 判断」のラベルがつき、文章や人の考え方に感情移入できる力などが求められていることが示され、大問の特徴の違いがラベルにおいても示された。

平成 25 年度及び平成 24 年度についても、「2. 解釈」のラベルが多く設問につけられた。3 年度に共通して付けられたラベルは、第 1 問では「2. 解釈」「19. 分解」、第 2 問では「2. 解釈」「11. 感情」「19. 分解」であった。「論理的な思考」6 ラベルで共通して付けられたものは無かった。

数学 I・数学 A

数学 I・数学 A についてまとめた結果を表 2 に示した。

「論理的な思考」6 ラベルについては、ほとんどの設問が「①理解・適用」であった。オーストラリアの 26 スキルについては、すべての設問が「1. 記憶」であった。「1. 記憶」の内容は「基本的な公式や法則、事柄などを記憶し必要に応じて思い出す力」である。両ラベルともに、数学 I・数学 A では、公式や法則を理解し、適切に適用する力が求められていることを示している。

平成 25 年度及び平成 24 年度についても、ほとんどの設問に「①理解・適用」及び「1. 記憶」のラベルがつけられた。これらのラベルに加えて、「論理的な思考」6 ラベルについては「④関係・洞察」が、オーストラリアの 26 スキルについては「3. 読む」「4. 変換」「26. 図式³」のラベルがつけられた設問もあった。

³26. 図式：スケッチ・作図・図式化する力

表 2: 外部評価者による評価のまとめ（平成 26 年度数学 I・数学 A）

	「論理的な思考」6 ラベル			オーストラリアの 26 スキル※		
	ラベル	設問数	配点	ラベル	設問数	配点
第 1 問 (8 問・20 点)	①理解・適用 なし	7 問 1 問	16 点 4 点	1. 記憶	8 問	20 点
第 2 問 (14 問・25 点)	①理解・適用	14 問	25 点	1. 記憶	14 問	25 点
第 3 問 (9 問・30 点)	①理解・適用	9 問	30 点	1. 記憶	9 問	30 点
第 4 問 (9 問・25 点)	①理解・適用	9 問	25 点	1. 記憶	9 問	25 点

※オーストラリアの 26 スキルの内容

1. 記憶：基本的な公式や法則，事柄などを記憶し必要に応じて思い出す力

数学 II・数学 B

数学 II・数学 B についてまとめた結果を表 3 に示した。

「論理的な思考」6 ラベルについては，多くの設問が「①理解・適用」であり，ほかに「②抽出・分析」「③把握・評価」のラベルがついた。数学 I・数学 A のように理解・適用するだけでなく，必要な情報を抽出したり，問題文を把握したりする応用を必要とすることが示された。

オーストラリアの 26 スキルについては，多くの設問が「1. 記憶」であり，ほかに「3. 読む」のラベルが第 5 問に，「4. 変換」「14. 推論」「15. 把握」のラベルが第 6 問についた。数学 II・数学 B では，第 1 問，第 2 問は必答問題であるが，第 3 問～第 6 問は 4 題から 2 題を選ぶ選択問題である。平成 26 年度の第 3 問は数列の問題，第 4 問はベクトルの問題，第 5 問は統計とコンピュータの問題（与えられた表や図を読む必要がある），第 6 問は数値計算とコンピュータの問題（プログラミングの問題）であり，選択問題の各大問の特徴が示される結果となった。

平成 25 年度及び平成 24 年度についても，平成 26 年度と同様に，ほとんどの設問に「①理解・適用」及び「1. 記憶」のラベルがつけられたほか，「②抽出・分析」「③把握・評価」，「4. 変換」「14. 推論」「15. 把握」のラベルがつけられた。これらのラベルに加えて「2. 解釈⁴」のラベルもつけられていた。

2.2 「2) 試験問題評価委員会報告書からの抜粋」の成果

試験問題評価委員会報告書からの抜粋では，センター試験の計 11 科目，3 年度分の本試験の各設問についての評価と各設問の正答率や項目相関（統計情報）を合わせ，一覧性が高くなるよう表の形式に整理し，資料を作成した。表の形式にまとめたことによって，各設問についての外部評価・作題方針・統計情報をすぐに見ることができるようになった。

2.3 「3) 作題部会部会長らによる試験問題総括」の成果

作題部会部会長らによる試験問題総括では，センター試験の問題作題部会の部会長らから計 51 件の回答を得た。作題部会においては，科目の各分野から多様な問題を出題するようにしている

⁴2. 解釈：言葉や他の記号（イラストなども含む）の意味を解釈する力

表 3: 外部評価者による評価のまとめ (平成 26 年度数学 II・数学 B)

	「論理的な思考」6 ラベル			オーストラリアの 26 スキル※		
	ラベル	設問数	配点	ラベル	設問数	配点
第 1 問 (19 問・30 点) 必答問題	①理解・適用	12 問	19 点	1. 記憶	16 問	26 点
	①理解・適用,②抽出・分析	6 問	9 点	なし	3 問	4 点
	①理解・適用,③把握・評価	1 問	2 点			
第 2 問 (11 問・30 点) 必答問題	①理解・適用	11 問	30 点	1. 記憶	10 問	27 点
				なし	1 問	3 点
第 3 問 (11 問・20 点) 選択問題	①理解・適用	6 問	11 点	1. 記憶	7 問	12 点
	①理解・適用,②抽出・分析	1 問	1 点	なし	4 問	8 点
	なし	4 問	8 点			
第 4 問 (13 問・20 点) 選択問題	①理解・適用	13 問	20 点	1. 記憶	13 問	20 点
第 5 問 (12 問・20 点) 選択問題	①理解・適用	7 問	12 点	1. 記憶	9 問	16 点
	①理解・適用,②抽出・分析	3 問	6 点	1. 記憶,3. 読む	1 問	2 点
	②抽出・分析	2 問	2 点	なし	2 問	2 点
第 6 問 (11 問・20 点) 選択問題	①理解・適用	6 問	11 点	1. 記憶	6 問	11 点
	①理解・適用,③把握・評価	4 問	8 点	1. 記憶,14. 推論,15. 把握	1 問	2 点
	②抽出・分析	1 問	1 点	14. 推論,15. 把握	3 問	6 点
			4. 変換	1 問	1 点	

※オーストラリアの 26 スキルの内容

1. 記憶：基本的な公式や法則，事柄などを記憶し必要に応じて思い出す力
3. 読む：表・図・地図・グラフが読めること
4. 変換：一つの表現形式を他の表現形式に変換する力
14. 推論：与えられた前提から結論を推論すること
15. 把握：部分的な情報から全体を把握できる力

こと、単なる知識を問う問題とならないよう、読解力や思考力を要求する問題と成るように常に意識して問題が作成されていることが分かった。

国語と数学について、試験問題総括の内容をまとめたものを次に示す。

国語

1. 論理的な文章である「評論」、文学的な文章である「小説」、古文、漢文と多様なジャンルにおける国語表現を素材に、(1回の試験でこれらが全て)出題されている。
2. 単なる文学史的な知識を問う暗記問題は出題しない一方で、漢字や慣用表現、文法等の日本語に関する知識は問う。これは、知識の活用力の基礎となるのは言語の運用能力であり、表現の方法を学ばなければ表現力は生まれず、思考力も判断力も形をなさないからである。
3. 基本的な語彙の把握と文章の理解によって「読解力」を測るとともに、内容を分析し自らの知識を拡充する「思考力」を測ることを目的に作問をしている。

1. に関し、センター試験に記述問題を導入する場合には出題数が限定される可能性があるため、高校生が身に付ける基本的なスキルを網羅的に、かつ系統的に出題できるような配慮が必要である。2. と 3. に関しては限られた制限の中で高度な思考力・判断力が測られるとしている。このために「作題時の工夫」として、素材文を読まずに受験テクニックで解けるような選択肢問題は作題しておらず、「条件付けられた複数の選択肢の文意や差異を素材文と照合しながら正確に把握し、選択する作業自体に高度な思考・判断力が求められている」としている。

数学

1. 「蓄えている知識を単に引き出しから取り出すだけの問題ではなく、広く思考力・判断力・応用力を要求する問題になるように、その問題を変形していくこと」を常に意識して問題作成している。つまり「型にはまらないように、僅かでも違いのある新傾向の問題を作り続ける」ことをしているという。
2. 高校3年までの学習内容を題材に、各分野から多様な問題を作成している。

1. により、「毎年新しい傾向の問題を出し続けることが知識の活用力を測っていることを担保する」のだと考えられる。2. によりセンター試験は「基盤となる計算力のみならず、代数的思考力、幾何学的思考力、解析的思考力などの数学的能力を測ることは、ひいては数学に限らない知識の活用力の判定に有効であろう」としている。

多くの作題部会長が指摘するように、センター試験は高等学校段階における数学の基礎的学習の達成度を適切に評価するため標準的な教科書で扱われる分野から偏りなく出題することが必要である。一方で、大学での学習において最も必要とされる資質の一つである数学的思考力・理解力を判定するために、高等学校学習指導要領・高等学校教科書の範囲を超えることなく、新傾向問題や融合問題を積極的に採り入れることが求められる。これらの目的を同時に達成するためには、また大量の試験データを公平に情報処理する必要があることから、従来では短時間の穴埋め式の問題という試験の形式がとられてきたものと考えられる。

なお、平成25年度センター試験問題総括には、「数学」の穴埋め式の問題と記述式問題の違いについての考察がある。

記述式問題の場合も、小問によって解答を誘導することが普通なので、誘導の有無が両者の主な相違点ではない。短い解答時間に多くの穴埋めをする問題では、解答の正否はその解答が指定枠に当てはまるか否かである程度判別できるので、後を振り返ることなく先へ計算を進めた方が高得点が期待できる。一方、長めの解答時間が与えられる記述式問題では、思考力を生かして、推論と計算の点検を十分行なうことが可能であり、そのように振り返りながらの解答方法の方が、結果として高得点が期待できる。今後、この性格の違いを十分認識して問題作成をしなければならない。

2.4 「4) センター試験の主要科目のまとめ」

センター試験の主要科目のまとめでは、それまでの評価結果を元に、センター試験に求められているものを整理し、センター試験の主要5科目で図っている内容の分析を行った。

また、センター試験の主要4科目の過去問の中から、特に思考力・判断力・表現力を測っていると評価された問題を良問例として選んだ。例を選んだのは検証事業aの研究担当者である。良問例には、実際の問題の画像（問題イメージ）に加え、良問と考えられるポイント（★）、測定される能力（思考力、判断力、読解力など、知識の活用力と関連する能力）及びその理由、正答率などを載せた。さらに試験問題評価員会報告書から、問題作成部会による各問題の出題意図を抜粋して「作題意図」の欄に、高等学校教科担当教員や教育研究団体による評価を抜粋して「評価コメント」の欄に載せ、1問を1枚にまとめた。ここでは、国語、数学（情報関係基礎）、物理、世界史について1問ずつ次に示す。

問題イメージ

平成26年度 本試験 第1問 問2

★枝葉末節ではなく文章の根幹を捉える力を要求している意図が明確であり、高校現場に指導の方向性を与えてくれる良問だと評価できる。

測定される能力

◎思考力、判断力
読解力
筆者の意図(なぜ中国に話を広げようとするのか)の理由を読み取らせる設問であるため。

正答選択肢: 4
得点率: 62%

- ④ となると、その過程で、ある特定の思考や感覚の型が形成されていたことにも、注意を向ける必要があります。といつても、忠や孝に代表される儒教道徳が漢文学習によって身についたと言いたいわけではありません。そうした側面がないとは言えないものでもなかったのです。
- ⑤ Aも少し広く考えてみましょう。
- ⑥ そもそも中国古典文は、特定の地域の特定の階層の人々によって担われた書物として始まりました。逆に言えば、その書きこぼしによって構成される世界に参入することが、すなわちその階層に属することになるわけです。どんなことばについてでもそうですが、B人がことばを得、ことばが人を得て、その世界は拡大します。前漢から魏晉にかけて、その書きこぼしは世界は古典世界としてのシステムを整えていき、高度なりテラシー(読み書き能力)によって社会に地位を(イ)シめる階層が、その世界を支えました。それが、士人もしくは士大夫と呼ばれる人々です。
- 問2 傍線部A「もう少し広く考えてみましょう」とあるが、それはなぜか。その説明として最も適当なものを、次の①～⑥のうちから一つ選べ。解答番号は⑥。
- ① 中国に目を転じて時代をさかのぼり、中国古典文に見られる思想と文学の共通点を考慮に入れることで、近世後期の日本において漢籍が知的世界の基礎になった根拠が把握できるから。
- ② 中国に目を転じて時代をさかのぼり、科挙を例に学問の制度化の歴史的起源に関する議論を展開することで、近世後期の日本において漢学が素養として公的に認知された理由が把握できるから。
- ③ 中国に目を転じて時代をさかのぼり、儒家だけでなく道家の思想も士大夫階級に受け入れられた状況を踏まえることで、近世後期の日本において漢文学習により知的世界が多様化した前提が把握できるから。
- ④ 中国に目を転じて時代をさかのぼり、中国古典文と士大夫階級の意識との関係を考察することで、近世後期の日本において漢文学習を通して思考や感覚の型が形成された過程が把握できるから。
- ⑤ 中国に目を転じて時代をさかのぼり、中国古典文に示された民情への視線を分析することで、近世後期の日本において漢学の専門家以外にも漢文学習が広まった背景が把握できるから。

作題意図

「言葉」と「人」の関係を通して、中国の士大夫が中国古典文を学ぶ過程でどのような現象が発生したのかを把握させる設問。中国古典文の本来的な性格の読み取りを修辭的に表現された内容把握を踏まえて確認させる意図を持っている。本文における言葉と人と世界との関係のありようを正確に理解しているか否かが正誤を分けたと考えられる。

評価コメント

筆者がなぜ中国に話を広げようとするのか、その理由を読み取らせる設問である。漢文を介して士大夫と士族が繋がっていることが理解できているかを確認する上で適切な設問である。
(高等学校教科担当教員)

問題イメージ

平成22年度 情報関係基礎 本試験

第2問 (必答問題)

次の文章を読み、下の問い(問1～3)に答えよ。(配点 50)

2以上の互いに異なる番号が書かれた札の束がある。ここから札を何枚か取り出し、番号が見えるように机の上に並べる。このとき、Aさん、Bさんの二人で並べた札の中から交互に札を取り合うゲームを考える。ルールは次のとおりである。

- ・ゲームはAさんが好きな札を1枚取ることから始まる。
- ・次に札を取るときには、直前に相手が取った札の番号と互いに素である番号の札を1枚取らなくてはならない(互いに素であるとは、1以外の公約数を持たないことである)。
- ・Aさん、Bさんが交互に札を取り、取ることでできる札がなくなるまで続ける。
- ・最後に札を取った方を勝つとする。

このようなゲーム設定のもとで、Aさんは、ゲームの作戦を考えることにした。Aさんが番号*i*の札を取り、Bさんが*i*と互いに素である番号*j*の札を取る。次に、Aさんが*j*と互いに素である番号*k*の札を取る。このようなゲーム経過をA(1)→B(1)→A(1)と記述する。

問1 次の文章の空欄ア～エに当てはまる数字をマークせよ。

取り出した札の番号が2, 4, 5, 6の場合、Aさんが最初に5の札を取ると、Bさんがどの札を取っても、Aさんは次に取る札がないので負け。ゲーム経過はA(5)→B(2), A(5)→B(4), A(5)→B(ア)のいずれかである。Aさんが最初に5以外の札を取れば、Bさんは5の札を取らざるを得ないのでAさんは必ず勝つ。例えば、Aさんが6の札を取ったときのゲーム経過は、A(6)→B(5)→A(4)またはA(6)→B(5)→A(イ)である。これらの場合には、すべての札が取られることなくゲームは終了する。取り出した札の番号が2, 3, 5, 6の場合、Bさんにうまく札を取られると、Aさんはどの札を取っても負けてしまう。そのゲーム経過の一つは

A(2)→B(3)→A(ウ)→B(エ)である。しかし、Bさんの札の取り方によっては、Aさんが勝てる場合もある。この場合のゲーム経過の一つは、A(2)→B(5)→A(6)である。

問2 次の文章の空欄オ～キに当てはまる数字をマークせよ。ただし、オ、カの解答の順序は問わない。

Aさんは、どの札が取れるかをわかりやすくするために、図を用いることを考えた。二つの札の番号が互いに素であるときのみ、それら二つの番号を線で結べば、次に取ることでできる札がすぐわかるようになる。

例えば、取り出した札の番号が2, 3, 5, 6の場合は図1のように表す。このように図で表したとき、ある札の次に取ることができるのは、その番号と線がつながっている番号の札だけである。図1でAさんが5の札を取った場合は、Bさんはそれとつながっている2, 3, 6のいずれかの札を取ることができる。一方、Aさんが2の札を取った場合は、Bさんはそれとつながっている3, 5のどちらかの札を取らないといけない。

次に、新たに2から9までの番号の札から何枚か取り出した。取り出した札の番号間の関係が図2(図の場合、取り出した札の番号は2, 3, オ, カ)であり、図2(イの場合、取り出した札の番号は2, 3, 4, 5, キ)である。

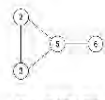


図1 番号間の関係

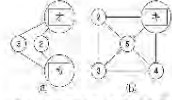


図2 取り出した札から得られた図

問3 次の文章の空欄ク～チに入れるのに最も適当なものを選び、次のページの解答欄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。また、ク、カ、キの解答の順序は問わない。

Aさんは、ゲームの作戦を立てやすくするために、図をさらに活用することを考えた。まず、図の中の線のいくつかを目印として太線にする。このとき、どの番号も2本以上の太線につながらないようにし、その上で可能な限り多くの太線にする。このような太線の引き方は、同じ札の集まりに対して何通りか考えられる。太線の引き方の例として、取り出した札の番号が2, 3, 5, 6, 14, 15, 21の場合を図3に、2, 3, 15, 21, 70, 140の場合を図4に示す。

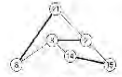


図3 すべての番号が太線につながっている場合

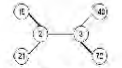


図4 太線につながらない番号がある場合

これからは、番号*i*を線が通る(行、A)と表す。図3では、すべての番号が太線(2, 3), (14, 15), (21)のいずれかにつながっている。一方、図4では、太線につながらない番号21, 140がある。図3の場合、Bさんにうまく札を取られるとAさんは絶対に勝てないことがわかる。例えば、最初にAさんが5の札を取ったとする。Bさんは、太線につながった地方のクの札を取ればよい。次にAさんがカの札を取ると、Bさんはやはり太線につながった地方の3の札を取ることができる。さらに、Aさんが14の札を取ると、Bさんが15の札を取り、Aさんは勝てない。このように、すべての番号が太線につながっているときには、太線につながった番号の札を最初に取った方は、相手にその太線の地方の番号の札を取られてしまうので、先に札を取った人は負けてしまうことがわかる。Bさんのこの作戦は、太線の選び方によらず成功する。図3の場合、新たな太線の組み合わせとして(ク), 21), (カ), 3), (キ), 15)を選

★2以上の番号が書かれた札の束から何枚かを取り出して並べ、二人が交互に札を取りあい直前に相手が取った札と互いに素である番号を取るゲームを考えると、並べる札の番号と枚数によってゲームの経過が異なることについて答えるアルゴリズムの問題。

正答率: 68%

測定される能力

◎思考力、読解力、知識の活用力
情報科学やオペレーションズリサーチでしばしば取り扱うゲームを題材とし、ネットワークグラフでの表現といった大学レベルの高度な内容まで組み入れているにもかかわらず、正しく読解すれば正解に至るから。

作題意図

情報技術に関する基本的事項として数理的ゲームを題材とし、論理的な思考能力と図形モデル化能力を問う問題とした。数理的ゲームの理解や最適解の探索を主に取り上げたが、単なる論理式問題にならず、受験者が理解しやすいように図形を利用して最適解を求める方法を題材にする形で出題している

評価コメント

読みこなすのが大変であるが、段階的に進めていける。「互いに素」の説明があり、良かった。「互いに素」を理解して、そのような番号を選び出すことができれば、後は問題文の指示どおり答えていくと、問1～3まですべて正解に至る。長文の日本語の問題文を読み進むだけの根気があれば、工業高等学校の生徒も解答できる、適切な問題だと思う。(日本教育工学協会)

物理 I (3/5)

問題イメージ

平成26年度 本試験 第3問 A 問2

第3問 次の文章(A・B)を読み、下の問い(問1～4)に答えよ。(配点 20)

A 媒質中を伝わる波動を、ベルトコンベアによる物品の搬送と対応させて、ドップラー効果を考えてみよう。

図1のように十分に長いベルトコンベアがあり、上に乗せたものを一定の速さ V で右に運んでいる。左側にいる作業員 A は、一定の速さ v_A で作業員 B に向かって移動しながら、一定の時間間隔 T_0 で小さな箱をベルトコンベアの上に乗せていく。右側の作業員 B は、運ばれてくる箱をベルトコンベアの端で回収する。ただし、 $v_A < V$ とする。

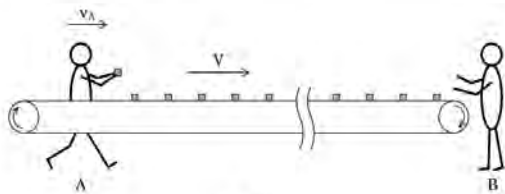


図 1

★媒質中を伝わる波動をベルトコンベアによる物品の搬送と対応させて、ドップラー効果を考察する設問

問2 次の文章中の空欄「ウ」～「オ」に入れる語の組合せとして最も適当なものを、下の①～⑥のうちから一つ選べ。 [13]

作業員 A、B をそれぞれ波源と観測者にみ立てて、ドップラー効果との対応を考えてみよう。箱の位置を波の山の位置、作業員 A が箱をおく時間間隔 T_0 を波源での波の周期、箱の速さ V を波の速さとみなすと、ベルトコンベア上にならぶ箱の間隔 d は観測される波の「ウ」、作業員 B が箱を受け取る時間間隔 T は観測される波の周期と解釈できる。

波源が運動してドップラー効果が起きているときは、波の「エ」は変わらず、「オ」が変化する。ベルトコンベアの搬送は波動とは異なる現象であるが、上記のように考えると、ドップラー効果を理解することができる。

	ウ	エ	オ
①	波長	波長	振動数
②	波長	速さ	振動数
③	振動数	波長	振動数
④	波長	振動数	速さ
⑤	振動数	振動数	速さ
⑥	振動数	速さ	波長

測定される能力

◎思考力, 判断力

基礎的な理解・知識を元にして、問題の設定に沿って適切に思考し、文意に沿う語を選択し、論理的につながりのあるような文章にする能力を問うから。

正答選択肢: 2 得点率: 70%

作題意図

問1の状況をドップラー効果に対応させたとき、波の波長、振動数、速さがどうなるかを問うた。問1の模型がどのように音波と対応するか、またどのような結果がドップラー効果と対応するかを問うた。このような対応が誤解を与えないようにベルトコンベアと波動とは物理として全く異なったものであることを問題中で述べた。

評価コメント

・ドップラー効果の本質的な理解を求めるように工夫されたすぐれた問いである。
(日本物理教育学会)

世界史B(1/5)

問題イメージ

平成24年度 本試験 第2問 B 問6

第2問 世界史上の国境について述べた次の文章A～Cを読み、下の問い(問1～9)に答えよ。(配点 25)

B 第一次世界大戦を境に、ヨーロッパの政治地図は大きく塗り替えられることとなった。中・東欧、バルト海地域、バルカン半島では、民族自決の原則を建前として③ポーランドをはじめ多くの国家が成立する一方、それまでの諸帝国は解体した。例えば④ドイツは、1919年6月に結んだヴェルサイユ条約によって、広大な領土を失った。こうして大戦後のヨーロッパでは、多くの新たな⑤国境線が引かれたのである。だが、これらの地域に曲がりなりにも適用された民族自決の理念は、他地域、とりわけアジア・アフリカにおいて直ちに実現されることはなかった。

問6 下線部⑤に関連して、国境の変更や領土の帰属をめぐる歴史について述べた文として波線部の正しいものを、次の①～④のうちから一つ選べ。 15

- ① アルザス・ロレーヌは、ウィーン会議の結果、ドイツ帝国領となった。
- ② 百年戦争の結果、イギリス王は、ポルドー以外の大陸の所領を失った。
- ③ テキサスは、アメリカ＝メキシコ戦争の結果、アメリカ領となった。
- ④ バルト3国は、第二次世界大戦が始まると、ソ連によって併合された。

★欧米において戦争によってどのように国境線が変更されたかを問う問題。

測定される能力

◎思考力、判断力、知識の活用

標準的な問題でありながら、「時間軸」と「空間軸」の両面を意識し、思考力や多角的な判断が必要とされるから。

正答選択肢:4 得点率:63%

作題意図

20世紀ヨーロッパにおいて第一次世界大戦を境にどのような国境線が引かれることになったのか、その際に掲げられた「民族自決」の自決とはどのようなものであったかについて考えることを意図して作問した。

評価コメント

・欧米で起きた戦争後の国境について、正文を選択する問題。3についてはやや細かい知識を必要とするが、1, 2, 4の判断は容易である。大問のテーマに即し、リード文を生かした良問である。「時間軸」と「空間軸」の両面が意識され、思考力や多角的な判断力を必要とする。(高等学校教科担当教員)

・国境の変更や領土の帰属に関する正文選択問題。いずれも標準的な問題である。(全国歴史教育研究協議会)

3 まとめ

センター試験は、大学（短期大学を含む）に入学を志願する者の高等学校段階における基礎的な学習の達成の程度を判定することを主たる目的とする試験である。しかし、入学試験であること、また、全国共通に実施される大規模な試験であることから、学習の達成程度を判定するという目的だけでなく、さまざまな要請に応える必要がある。

本検証事業の結果、センター試験は、マークシート式でも一定の思考力等を測ることができるように歴代の作題委員の創意工夫によって多くの洗練を重ねてきていることが確認された。特に、知識の習得状況の評価に優れていることに加えて、解答にマークシートを用いながらも、与えられた問題を分析的に思考・判断する能力も測定する問題であると評価されていることが確認された。

一方で、複数の情報を統合し構造化して新しい考えをまとめる思考・判断の能力や、その過程や結果を表現する力などを評価することができるよう、一層の改善が求められる。

引用文献

国立教育政策研究所 (2013). 特定の課題に関する調査（論理的な思考）調査結果～21世紀グローバル社会における論理的に思考する力の育成を目指して～. http://www.nier.go.jp/kaihatsu/tokutei_ronri/pdf/10_tyousakekka.pdf

大学入試センター (2014). 平成 26 年度大学入試センター試験試験問題評価委員会報告書。

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.

参考資料

「平成 26 年度大学入試センター試験試験問題評価委員会報告書」（大学入試センター，2014）から今後の留意点や要望についてまとめてある部分を引用し、今後の課題を示す。国語と数学を取り上げる。

国語

第 1 問及び第 2 問について「今後の問題作成に当たっての留意点又はまとめ」で述べられている内容は次の通りである。

（第 1 問に関して）

- 評論問題に関しては、今後とも、高等学校の生徒に日常的になじみやすい話題や視点ばかりでなく、その論旨が文章全体から明快に読み取れる文章を選定することが望ましい。
- 各設問においては、選択肢の文言を精査・精練して、内容上は適切にして、かつ晦渋⁵に過ぎぬ平明な表現を盛り込んでいく必要がある。
- 全体と部分の読み取りについても、傍線を付す箇所を工夫を凝らすなどして、過度の負担が受験者に生じないように心掛けたい。

（第 2 問に関して）

- 小説問題については、本年度は戦時下の作品を取り上げ、作品の全文を掲載した。分量は多かったが、人物の心情や展開は読み取りやすく、表現上の特徴も指摘しやすい作品であった。今後も受験者にとって有益な内容の作品を題材として選び、小説読解の基本的能力を見る設問を心掛けたい。
- 指摘されている素材文の分量の多さについては、受験者の負担にならないように留意したい。

⁵かいじゅう：ことば・文章などが難しく、意味がわかりにくいこと。（明鏡国語辞典より）

部分を問うか全体を問うかに関しては、第1問の問2について「誘導・ヒントのための問題と考えると、箇所としてはいい。しかし全体を見渡さなければならない問いになっており、不適切。『誘導としての問2』であれば、部分的な読みを求める問いでいいのではないか。(日本国語教育学会)」との指摘があった。一方、高等学校教科担当教員の意見・評価では「傍線部前後の情報処理だけでは正解できない設問を多く設定したことは、読解力や思考力を身につける重要性や意義を示したものとして高く評価したい」とあり、問題作成部会の見解では「その場合、事によっては、問題文全体を適切に把握せずに、選択肢から答えを選んでいくという受験技術的弊害を助長しかねないことが懸念される。」と述べられている。各設問で問う内容について、今後も議論を続けていく必要がある。

素材文については、第1問については「扱っている内容が受験生の生活感覚から離れていた(日本国語教育学会)」, 第2問については「時代背景などの前提知識を必要とする文章だった(日本国語教育学会)」との指摘がなされていた。また、第2問の素材文の文字数が多かったことから「各設問と本文の照合に時間の掛かった受験者は多かったはずである(高等学校教科担当教員の意見・評価)」との指摘もあった。素材文は、国語の問題の質を決定する要と考えられる。素材文の内容、表現、分量などについて、どのような素材文が適切であるか、議論を続けていく必要があるであろう。

また、現代文ではないが、第3問(古文)の問5に関しては「主語で誤答と判断できるのは③だけである。後は三つの応答の的確な理解を必要としているため、3行にわたる選択肢を吟味するのは相当厳しかったと思われる。また、消去法で正解が①であると判断できるものの、『子どもたちのことはあととよろしくと言っている』という表現が適当かどうか疑問が残る。(高等学校教科担当教員の意見・評価)」との指摘があった。読解力を求める問題において、消去法で正解に至る設問が妥当かどうか、また選択肢の表現が適当か否か、等の選択式特有の課題についても検討する必要があるであろう。

数学

数学Ⅰ・数学A及び数学Ⅱ・数学Bについて「高等学校教科担当教員の意見・評価」にまとめられている今後のセンター試験への要望等は次の通りである。

(数学Ⅰ・数学Aに関して)

- 今後とも高等学校学習指導要領に定められた範囲内の内容であるとともに、知識・理解、表現・処理だけでなく、数学的な見方や考え方なども十分評価できるような出題をお願いしたい。
- 今後とも問題の精選や数値の工夫など受験者が考える時間の確保に加え、問題冊子の余白への配慮についてお願いしたい。
- 「数学Ⅰ」と「数学Ⅰ・数学A」間について、平均点の差をなくす配慮よりも難易度の差異がないような工夫の継続をお願いしたい。また、平成27年度においては、新旧課程間の難易度の差異がないような工夫をお願いしたい。
- 本試験と追・再試験間について、難易度の差が適切であるような工夫の継続をお願いしたい。

(数学Ⅱ・数学Bに関して)

- 基礎的・基本的な事項の理解を問う問題と、数学的な思考力・計算力等を問う問題がおおむねバランス良く出題されていた。来年度以降においても、この出題傾向を継続していただきたい。
- 「数学Ⅱ・数学B」の出題分野が特定分野に固定されないように、出題分野の融合も考えるなど、今後ともより工夫、配慮した出題を考えていただきたい。
- 受験者が、実際に問題冊子の余白に図を書いたり、計算することを考慮して、より一層、問題文のレイアウトの工夫や、冊子のサイズを変更するなど、余白を作り、計算ができる部分を増やすようにしていただきたい。
- 60分という時間の制約の中で受験者が解答することを考えて、分量と難易度のバランスをとり、今後も平均点を目標とする60点に近づけていただきたい。
- 適切な誘導があり、先の解答の見通しが立つような問題文の表現が工夫され、基礎的な知識から数学的な思考力や応用力まで問うものとなっており、評価できる。

誘導に関しては、数学Ⅱ・数学Bにおいて日本数学教育学会から次のような指摘がある。「全体的に受験者が本来発想するべきところが誘導されており、受験者の数学的な思考力を判断する出題になりにくくなっている。ただし、適切な誘導によって難易度の調整がきちんとされているとも考えられる」。誘導が適切かどうか、数学的な思考力等を問うことができるかどうかについて、高等学校教科担当教員と日本数

学教育学会とで異なる意見となっている。誘導によって難易度が調整され、幅広い層の学力を測ることが可能となっている側面があるが、誘導の是非を含め、どのような誘導が適切であるかなど、誘導に関する議論を進める必要があるであろう。

b 新テスト試作問題の作成と評価

1 「b. 新テスト試作問題の作成と評価」の概要

「b-1. 作業グループによる新テスト試作問題の作成」および「b-2 試験問題評価のためのモニター調査の実施」では、中教審答申「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」（平成26年12月22日）において示された「大学入学希望者学力評価テスト（仮称）」の実現に向けて、専門家会議（「高大接続システム改革会議」）の議論を踏まえつつ問題イメージの作成を行うとともに、それらの問題の信頼性・妥当性についての検討を行った。

「b-1. 作業グループによる新テスト試作問題の作成」では、新テストWGで示された試験実施の方針および試験問題作成の枠組みに基づき、特に新たな特徴を持つ科目について、共通試験の問題作成経験のある大学教員等で構成する作業グループを試験科目ごとに構成して、サンプル問題の作成を行った。作業グループでは、該当科目の試験の作題方針の決定、サンプル問題（案）の作成を行った。「大学入学希望者学力評価テスト（仮称）作業チーム」として、「国語」「数学」「地理歴史（世界史）」「理科（物理）」の4部会を設置した。試験問題作成経験のある大学教員等各5名で構成した各部会において、該当科目の試験の作題方針の決定、サンプル問題（問題イメージ）の作成を行った。また、上記作業チームとは別に、「英語」についても、専門業者2社に依頼し、サンプル問題（案）の作成を行った。

「b-2 試験問題評価のためのモニター調査の実施」では、b-1によって作成された試作問題の評価のためのモニター調査を実施した。作業グループへの早い段階でのフィードバックを行うために、実施規模は数百人とどめたが、モニター参加者へのアンケートや、センター試験得点との比較など付随情報の取得も行った。

b-2のモニター調査で得たデータを分析して、サンプル問題の統計的特性について検討を行った。

2 各作業チームによる問題イメージの検討

図1に、作業チームによる作業経過を示す。

5月21日に開催された全体会（国語、数学、地理歴史（世界史）、理科（物理））において各部会に分かれて部会長と副部会長を決定した。以降、各部会ごとに検討を行った。現行のセンター試験で測れている能力、測定が不十分だった能力について、各部会で検討すると共に、新テストでどのような能力を測定するべきか、という作題方針を検討した。さらに、測定しようとする能力を測るための試作問題とその作題意図を、問題イメージ報告書としてまとめる作業を行った。各部会がまとめた作題方針と問題イメージ報告書は、7月1日付で文部科学省に報告した。それらの作業と並行して、モニター調査で用いる問題についての検討も行った。

9月17日に作業チームの部会長会議が開催され、次のような作業内容を確認した。

- (1) 『『大学入学希望者学力評価テスト（仮称）』の作問イメージ』の作成：所定の様式に合わせて、これまでのセンター試験の現状、現状を踏まえた今後の対応、テストによって評価しようとする力をまとめること。9月15日の高大接続システム改革会議で示された「基本的枠組み」とテストによって評価しようとする力の対応を記入すること。
- (2) 7月1日付で報告した問題イメージ報告書の見直し：問題発見・解決のプロセスの中で働く思考力・判断力・表現力等の観点に留意して見直す。
- (3) 新たに問題イメージを追加で作成。

(1)に対応して、「作問方針イメージ」への意見を、10月8日に文部科学省に報告した（2.4節）。(2)と(3)に対応した「問題イメージ報告書」は、10月30日付で文部科学省に報告された。国語と数学については、10月30日付で報告した「問題イメージ報告書」をもとに、文部科学省が手を加えたものが、12月22日の高大接続システム改革会議で公開された。

これらの作業と並行して、第1回モニター調査で用いる問題の選定、および、ブラッシュアップも行った。また、10月30日付で「問題イメージ報告書」を文部科学省に提出した後は、第2回モニター調査に向けた準備を行った。また、部会では、問題の採点に関する検討および採点作業も行った。

モニター調査は、サンプル問題（問題イメージ）の解答データを収集して、問題の特徴を検討するために、実施された。

さらに、1月下旬には、地理歴史（世界史）と理科（物理）の部会に対して、「多肢選択式」の問題イメージを準備するよう文科省から要請があり、2月15日付で報告された。この報告をもとに、2月17日の高大接続システム改革会議において『『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マーク式問題イメージ例』が公表された。

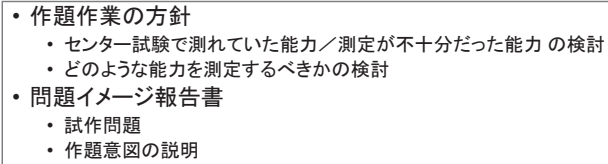
表1～4に、各作業部会の開催日時と作業内容を示す。部会開催日以外にも、メールのやりとりによる検討が行われた。また、部会開催日以外に採点作業を行った部会もある。

各部会のオブザーバーとして、国語は椎名、数学は桜井、地理歴史（世界史）は荒井、理科（物理）は伊藤が同席して、議事録の作成にあたりと共に、該当科目のモニター調査の準備や分析を行った。大津研究開発部長は、やむを得ない場合を除いて、ほぼすべての作業部会に同席した。

作業経過

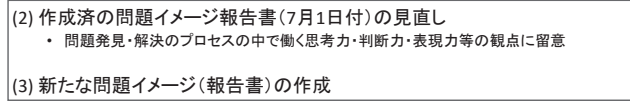
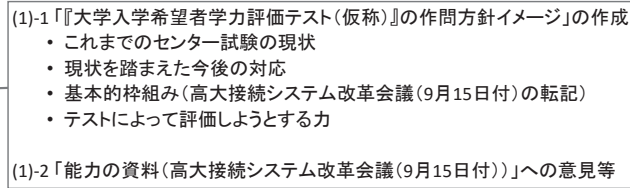
- 4月～5月中旬:作業チームの立ち上げ
 - 国語／数学／理科(物理)／地理歴史(世界史)

- 5月21日:作業チームによる検討を開始



- 7月1日:各部会での検討案を文科省入試室に報告
- 7月～:各作業チームにおいて、試作問題のブラッシュアップ
モニター調査に向けた作業

- 9月17日:部会長会議にて文科省からの依頼



10月1日に入試室に報告

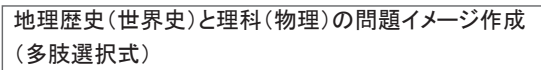
10月6日に入試室から修正要求

10月8日に修正版を報告

10月30日に入試室に報告

- 12月19日:部会長会議にて文科省から説明
12月22日の高大接続システム改革会議で配付予定の資料についての説明

- 1月下旬:文科省から依頼



2月15日に入試室に報告

図 1: 部会における作業経過

表 1: 国語部会の開催日時と作業内容

開催日	開催時間	作業内容
6/7(日)	10～12時	<ul style="list-style-type: none"> ● 部会長・副部会長の確認 ● 問題イメージについての議論 ● 連動型複数選択問題方式についての議論 ● 問題構成の枠組みについての議論 ● 各委員の持ち寄り資料に基づく議論 ● 作題方針についての議論 ● 次回までの作業分担の決定
6/18(木)	14～17時	<ul style="list-style-type: none"> ● 問題イメージの作成に向けての議論 ● 問題イメージ報告書の書式の検討 ● 持ち寄った問題の検討 ● 問題イメージとしてブラッシュアップする問題の選定 ● モニター調査の目的の確認
6/26(金)	10～12時	<ul style="list-style-type: none"> ● 問題イメージの示し方についての検討 ● 連動型複数選択問題についての検討 ● 記述問題の採点についての検討 ● 各委員が作成した問題案のブラッシュアップ ● モニター調査の実施計画の検討
8/25(火)	13～17時	<ul style="list-style-type: none"> ● 文科省からの要望への対応についての検討 ● 問題の見直しについての検討 ● モニター調査についての検討
9/16(水)	13～17時	<ul style="list-style-type: none"> ● 問題の追加作成についての検討 ● 第1回モニター調査の目的についての検討 ● 問題案の検討 ● 第1回モニター調査で用いる問題の決定
9/27(日)	10～15時	<ul style="list-style-type: none"> ● 「作問方針のイメージ」についての検討 ● 「能力の資料」についての検討 ● 第1回モニター調査で使用する問題案の検討
10/1(木)	13～16時	<ul style="list-style-type: none"> ● 「作問方針のイメージ」の確認 ● 第1回モニター調査で使用する問題案の検討 ● 第1回モニター調査の問題冊子の構成の検討
10/12(月)	13～16時	<ul style="list-style-type: none"> ● 第1回モニター調査で使用する問題についての検討 ● 文科省提出用の問題作成方針についての検討 ● 第1回モニター調査の採点方法についての検討

(次ページに続く)

(前ページからの続き)

開催日	開催時間	作業内容
12/19(土)	13～16時	<ul style="list-style-type: none"> ● 「作問方針のイメージ」の確認 ● 第2回モニター調査で使用する問題についての検討 ● 第2回モニター調査の問題の採点作業についての検討 ● オブザーバーより、第1回モニター調査の分析結果の報告

表 2: 数学部会の開催日時と作業内容

開催日	開催時間	作業内容
6/5(木)	10:00～12:00	<ul style="list-style-type: none"> ● 新テストに関する論点の整理 ● 作題方針についての議論 ● 問題イメージについての議論
6/19(金)	10:00～12:00	<ul style="list-style-type: none"> ● 測定しようとする能力の枠組みと分類の検討 ● 問題イメージについての議論 ● 文科省からの要望への対応についての検討
6/30(火)	10:00～12:30	<ul style="list-style-type: none"> ● 問題イメージ報告書、対応する問題イメージの検討 ● 数学部会の作業方針の検討 ● 作題方針についての議論 ● 解答形式についての議論 ● 文科省からの要望への対応についての検討
7/31(金)	13:00～17:00	<ul style="list-style-type: none"> ● 作題方針についての議論 ● モニター調査を行う問題イメージ案の検討
9/1(火)	13:30～17:30	<ul style="list-style-type: none"> ● モニター調査の目的、解答形式の検討 ● 第1回モニター調査で使用する問題と解答形式の検討 ● 文科省からの要望への対応についての検討
9/18(金)	10:00～15:00	<ul style="list-style-type: none"> ● 第1回モニター調査で使用する問題（問題冊子）、解答用紙、アンケート用紙の決定 ● 記述式設問の採点方法の検討 ● 正解表・採点指示書の作成 ● 文科省からの要望への対応についての検討
10/23(金)	10:00～14:00	<ul style="list-style-type: none"> ● 記述式設問の採点方法の検討 ● 正解表・採点指示書（および補足資料）の修正 ● 作題方針についての議論 ● 問題イメージについての議論（第2回モニター調査の準備） ● 文科省からの要望への対応についての検討 ● オブザーバーより、モニター調査答案の採点体制、データ入力体制の報告

(次ページに続く)

(前ページからの続き)

開催日	開催時間	作業内容
11/16(月)	10:00～14:15	<ul style="list-style-type: none"> ● 第1回モニター調査(10月)の採点結果に基づく採点基準の見直し ● 第2回モニター調査の目的, 解答形式の検討 ● 第2回モニター調査で使用する問題の検討 ● 文科省からの要望への対応についての検討 ● オブザーバーより, 第1回モニター調査(10月)の実施状況, 採点結果・アンケート集計結果(速報)の報告
12/7(月)	10:30～15:30	<ul style="list-style-type: none"> ● 公開予定の問題イメージについての議論 ● 第1回モニター調査(10・11月)の採点結果に基づく採点基準の見直し ● 第2回モニター調査の実施方法についての議論 ● 第2回モニター調査(1月)で使用する問題(問題冊子), 解答用紙, アンケート用紙の決定 ● 記述式設問の採点方法の検討 ● 正解表・採点指示書の作成 ● 文科省からの要望への対応についての検討
1/25(月)	10:00～15:00	<ul style="list-style-type: none"> ● 第2回モニター調査答案(記述式問題)の採点作業
2/4(木)	10:00～15:15	<ul style="list-style-type: none"> ● 第2回モニター調査答案(記述式問題)の採点作業(続き) ● 第2回モニター調査で出題した記述式問題の採点作業についての議論 ● 正解表・採点指示書の修正 ● 問題訂正の作成 ● 公開された問題イメージの採点基準の検討 ● 報告書の作成準備 ● オブザーバーより, 第2回モニター調査の実施状況の報告

表 3: 地理歴史(世界史)部会の開催日時と作業内容

開催日	開催時間	作業内容
6/5(金)	14～18時	<ul style="list-style-type: none"> ● 部会長・副部会長の確認 ● 連動型複数選択問題方式についての説明 ● 各委員の持ち寄り資料に基づく議論 ● 問題イメージについての議論 ● 出題範囲や問題の難易度に関する議論

(次ページに続く)

(前ページからの続き)

開催日	開催時間	作業内容
6/19(金)	14～18時	<ul style="list-style-type: none"> ● 問題イメージ報告書についての議論 ● 記述式の導入に関する議論 ● 各委員が作成した問題案の検討
8/5(水)	14～18時	<ul style="list-style-type: none"> ● 第1回モニター調査の目的についての検討 ● 第1回モニター調査で用いる問題についての検討 ● 第1回モニター調査で用いるアンケートについての検討
9/6(日)	13～17時	<ul style="list-style-type: none"> ● 「作問方針のイメージ」についての検討 ● 問題イメージ報告書の修正についての検討 ● 問題イメージの作成についての検討 ● 第1回モニター調査で用いる問題についての検討 ● 第1回モニター調査で用いるアンケートについての検討
11/29(日)	13～17時	<ul style="list-style-type: none"> ● 提出した問題イメージについての確認 ● オブザーバーより、第1回モニター調査の実施結果の概要(速報値)の報告 ● 第2回モニター調査で用いる問題についての検討 ● 第2回モニター調査で用いるアンケートについての検討
12/8(火)	17～20時	<ul style="list-style-type: none"> ● 問題イメージについての再検討 ● 第2回モニター調査で用いる問題についての検討 ● 第2回モニター調査で用いるアンケートについての検討
12/24(木)	15～18時	<ul style="list-style-type: none"> ● 問題イメージについての再検討 ● 第2回モニター調査の採点方法などの検討
2/13(土)	14～18時	<ul style="list-style-type: none"> ● 問題イメージ(多肢選択式)の検討 ● オブザーバーより、第2回モニター調査の実施結果の概要(速報値)の報告

表 4: 理科(物理)部会の開催日時と作業内容

開催日	開催時間	作業内容
6/16(火)	10:00～17:30	<ul style="list-style-type: none"> ● 新テストに関する論点の説明と議論 ● CBT についての説明と議論 ● 問題イメージについての議論 ● 連動型複数選択問題方式の説明
6/30(火)	10:00～17:30	<ul style="list-style-type: none"> ● 問題イメージについての議論 ● 第1回モニター調査に向けての検討事項等の確認 ● 作業チーム委員構成、新テストWGとの関係の確認 ● 作業チーム理科部会の作業方針の作成

(次ページに続く)

(前ページからの続き)

開催日	開催時間	作業内容
7/14(火)	10:00～17:30	<ul style="list-style-type: none">● 高大接続システム改革会議および新テストWGの現状報告と議論● 問題イメージおよび評価する能力と評価方法についての議論● 第1回モニター調査に関する検討
8/4(火)	10:00～17:30	<ul style="list-style-type: none">● 新テストWGの議論内容と作成中の問題イメージとの比較検討● 中教審答申の確認● 解答形式(数式記述式, 文章記述式, 選択式)の検討● 新テストとしての理科の実施方法についての検討● 連動型複数選択問題方式の理科分野への適用可能性の検討● 思考力・判断力・表現力に関わる各種調査の問題例の確認
8/26(火)	10:00～17:30	<ul style="list-style-type: none">● 改革会議の内容報告と今後の作業についての説明(文科省専門官)● 問題イメージについての議論● 問題内容と解答用紙の対応の検討
9/17(木)	10:00～17:30	<ul style="list-style-type: none">● モニター調査の理科領域受験者に求められる属性の検討● モニター調査の受験者募集時に聴取する受験者属性の確認● 作業チーム部会長会議の内容についての報告● 文科省からの要望への対応についての検討
9/30(水)	10:00～17:30	<ul style="list-style-type: none">● 改革会議「中間まとめ」の「能力の資料」への意見等の検討● 第1回モニター調査の問題冊子デザインの確認● 「作問方針のイメージ(理科(物理))」の検討
10/14(水)	10:00～17:30	<ul style="list-style-type: none">● 「作問方針のイメージ(理科(物理))」の修正● 問題イメージ(報告書)の見直し
10/23(金)	10:00～13:00	<ul style="list-style-type: none">● 「作問方針のイメージ(理科(物理))」改定案の検討● 新たな問題イメージの検討● 第1回モニター調査の解答用紙デザインの検討
11/11(水)	10:00～17:30	<ul style="list-style-type: none">● 新テストの実施方法の見通しについての確認● 記述式の答案の採点方法の現状と課題についての議論● 第1回モニター調査(先行実施分)の答案内容の確認● 第1回モニター調査の記述式問題の採点基準の検討
11/18(水)	10:00～17:30	<ul style="list-style-type: none">● 第1回モニター調査の記述式問題の採点基準の検討● 第1回モニター調査の記述式問題の採点指示書の作成

(次ページに続く)

(前ページからの続き)

開催日	開催時間	作業内容
12/1(火)	10:00～17:30	<ul style="list-style-type: none">● 改革会議における「記述式導入の考え方について」の確認● 第1回モニター調査の記述式問題の採点指示書の確定● 第2回モニター調査に向けての検討事項の確認● 第2回モニター調査の問題内容の検討
12/8(火)	10:00～17:30	<ul style="list-style-type: none">● 第2回モニター調査の問題内容の検討
12/18(金)	10:00～14:00	<ul style="list-style-type: none">● 新テストの議論の現状および今後の方向性の説明（文科省）● 第2回モニター調査の問題内容の検討
1/13(水)	10:00～17:30	<ul style="list-style-type: none">● 高大接続システム改革会議の現状報告と議論● 問題イメージの検討
2/1(月)	10:00～17:30	<ul style="list-style-type: none">● 理科領域の問題イメージの公開に向けての検討● 新テストで評価すべき内容とマーク式問題イメージの検討
2/29(月)	10:00～17:30	<ul style="list-style-type: none">● 高大接続システム改革会議・新テストWGについての報告● 第1回モニター調査記述式答案の採点作業状況の経過報告● 第2回モニター調査の答案確認● 新たに導入した記述問題の採点基準の検討

3. 英語 4 技能を測定する問題イメージの作成

3.1 はじめに

中央教育審議会答申「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」（平成 26 年 12 月 22 日）において、現在の大学入試センター試験に代わり「大学入学希望者学力評価テスト（仮称）」（以下「学力評価テスト」）を導入することが提言された。英語について同答申では、以下のような提言が出されている。

特に英語については、四技能を総合的に評価できる問題の出題（例えば記述式問題など）や民間の資格・検定試験の活用により、「読む」「聞く」だけでなく「書く」「話す」も含めた英語の能力をバランスよく評価する。また、他の教科・科目や「合教科・科目型」「総合型」についても、英語についての検討状況も踏まえつつ、民間の資格・検定試験の開発・活用も見据えた検討を行う。

さらにグローバル化の進展の中で、国際共通語である英語力の向上が日本の将来にとって重要な課題であるという認識のもと、初等中等教育を中心とした「英語教育の在り方に関する有識者会議」が開催されてきた。「英語教育の在り方に関する有識者会議」報告書（平成 26 年 9 月 26 日）では、グローバル化に対応した英語教育改革の五つの提言がなされている。そのうちの一つ、「改革 3. 高等学校・大学の英語力の評価及び入学者選抜の改善」を表 1 に示す。

表 1 高等学校・大学の英語力の評価及び入学者選抜の改善
（「英語教育の在り方に関する有識者会議」報告書）

-
- 生徒の 4 技能の英語力・学習状況の調査・分析を行い、その結果を、教員の指導改善や生徒の英語力の向上に生かす。
 - 入学者選抜における英語力の測定は、4 技能のコミュニケーション能力が適切に評価されることが必要。
 - 各大学等のアドミッション・ポリシーとの整合性を図ることを前提に、入学者選抜に、4 技能を測定する資格・検定試験の更なる活用を促進。
そのため、学校、テスト理論等の専門家、資格・検定試験の関係団体等からなる協議会を設置し、
 - ・適切な資格・検定試験の情報提供、
 - ・指針づくり（学習指導要領との関係、評価の妥当性、換算方法、受験料・場所、適正/公正な実施体制等）、
 - ・試験間の検証、英語問題の調査・分析・情報提供等の取組を早急に進めることが必要。
 - 「達成度テスト」の具体的な検討を行う際には、連絡協議会の取組を参考に英語の資格・検定試験の活用の在り方も含め検討。
-

「高大接続改革実行プラン」(平成 27 年 1 月 16 日文部科学大臣決定)に基づき、現在、「高大接続システム改革会議」等において高大接続改革の実現に向けた具体的な方策について検討が行なわれている。英語については、①従来のセンター試験でも実施されてきた「聞くこと」と「読むこと」に、「書くこと」と「話すこと」を加えた 4 技能の評価、②実施にあたり、4 技能試験の経験が蓄積されている、民間の資格や検定試験の更なる活用を目指すことが期待されている。

そこで、民間事業者等の知見も得つつ、より多様な観点から検討材料を得た上で問題イメージ等を作成することにした。

3.2 問題イメージ提出までの経過

問題イメージ提出までの経過を表 2 に示す。

表 2 実施の経過

年月日	事項
平成27年 11月17日	公募開始
11月30日	参加表明書提出期限
12月 3日	提案書提出期限
12月 7日	総合評価委員会開催
12月 9日	評価シート提出期限
12月 9日以降	契約書締結
12月28日	問題イメージ原案提出
平成28年 1月 9日以降	中間報告書の提出
3月14日	最終報告書の提出(問題イメージの提出)

3.3 問題イメージの要件

入札に際しての要求要件は、以下の通りである。

3.3.1 試験問題作成の実績

(試験問題作成の実績)

- ・ 過去に類似の英語の試験問題作成業務を行なった実績を有すること。

3.3.2 対象・実施回数

(対象・実施回数)

- ・ 年間複数回実施し、総数は 50 万人を想定している。実施回数及び 1 回当たりの受験者数は競争加入者が想定すること。

3.3.3 出題方法・内容

(問題イメージ作成対象科目等)

ア 問題イメージ作成対象科目は英語とし、出題範囲は、現行の高等学校学習指導要領(平成 21 年 3 月告示)を踏まえ、英語の四技能(「聞くこと」、「話すこと」、「読むこと」、「書くこと」)を測るものとし、英語の四技能及び技能統合型の測定が可能なものであること。

(問題内容)

イ 四技能(「聞くこと」、「話すこと」、「読むこと」、「書くこと」)及び技能統合型とし、高大接続システム改革会議「中間まとめ」(平成 27 年 9 月 15 日)(以下、「中間まとめ」)に示されている提言及び英語に関する提言を踏まえた内容とすること。

※参考：Ⅲ 高大接続システム改革実現のための具体的方策

3. 大学入学者選抜改革 (2)大学入学希望者学力評価テスト(仮称)の導入 ウ② 思考力・判断力・表現力を構成する諸能力をより適切に評価するため、多肢選択式の問題に加え、問題に取り組むプロセスにも解答者の判断を要する部分が含まれる問題、記述式の問題等。

(構成、試験時間)

ウ 1 回の試験時間は「聞くこと」、「読むこと」、「書くこと」の三技能及び技能統合型合計で 110 分程度、別に「話すこと」10 分程度のテスト構成・時間を想定し、適切な問題数及び試験問題の全体構成を検討の上、問題イメージを作成する。

(問題イメージの作成方法及び公表)

エ 問題イメージは、既存のものではなく、本事業において新たに作成するものとする。ただし、過去に使用した問題を用いることに合理的な理由がある場合にはセンターと協議すること。また問題イメージは、センターと調整の上、作成すること。(なお、作成した問題イメージは、センター等において精査した上で、公表を予定している。また必要に応じ問題イメージ原案の作成者を公表することがある。)

(出題数)

オ 四技能及び技能統合型の大き問を柱に全体構成を提案するとともに、それぞれ 2 問程度以上を作成すること(10 問～20 問程度)。

(出題形式)

カ 選択式を中心に、記述式も一部含めること。

(C B T を活用した問題)

キ C B T (Computer Based Testing) により出題する問題を含めるものとし、その特性を活かした問題(例：動画等を活用した問題、連問式※、の問題など)も一部含め、その一部の問題でプレゼンテーションとして実演できること。

※出題する課題に対し、複数の問題を順次出題し、解答を求める方式(一度解答し、次の問題に進んだ場合は前の問題に戻れない仕組み)

(出題の難易度)

ク 作成する各問題は、「外国語の学習、教授、評価のためのヨーロッパ共通参照枠」(CEFR)の A 1 上位層から B 1 層を中心として測定が可能であり、CEFR におけるレベルとの関連付けをすること。言語使用場面：買物や旅行などの特有の表現が使われる場面、家庭での生活や学校での学習や活動などの生徒の身近な暮らしや社会での暮らしにかかわる場面、本や新聞などの多様な手段を通して情報などを得る場面等

(正答(例)の作成)

ケ 問題イメージの正答(例)を作成すること。

(採点方法等の作成)

コ 採点方法、採点基準の案等を作成すること。

(成績等の結果の表示方法について)

サ 成績等の結果の表示方法を示すこと。

(問題イメージ報告書の作成)

シ 試験問題の全体構成とともに、各問題の①問題概要(特徴)、②出題意図、③測定する能力・資質、「中間まとめ」別添資料5の内容との対応関係、④解答形式(多肢選択、短文等)、⑤実施方法(紙筆式、C B T等)、⑥CEFRにおけるレベル、⑦配点、⑧問題に対応する学習指導要領の項目を問題ごとに「問題イメージ報告書(中間報告書)」として作成し契約締結日の翌月の同日までに提出すること。中間報告書を踏まえ、平成28年3月14日までに「最終報告書」を提出すること。「最終報告書」の内容は別途センターと協議すること。

3.3.4 その他

(記述式試験、C B T、等化等の実施方法の提案)

ア 実施回数や想定される受験者数及び問題に関連して記述式試験のあり方、C B T、等化(I R Tを含む。)などの実施方法等に関する提案を行うこと。

(教育上の配慮)

イ 時事的な社会現象、特定地域に係る事項、思想、信条、宗教、民族、性などに関する事項については、教育的に公正であるよう配慮すること。

(問題イメージ等の提出時期)

ウ 平成27年12月中に問題イメージ原案等を提出し、必要な調整をした上で平成28年3月までに問題イメージ等を提出すること。センターにおいては、問題イメージ原案提出時には十分確認の上、必要に応じて修正等を行う。

(出題者)

エ 出題者は英語四技能に関する国内外のテスト開発・作成に従事した経験を有する者を複数名選定するものとし、略歴及び業務実績がわかる書類を添付すること。

(提案内容の具体性)

オ 仕様書記載の内容が提案書に具体的な表現で提案されていること。

(実施方法の妥当性、実現性)

カ 実施方法に妥当性があり、実現可能なものとなっていること。

3.4 問題の構成と内容

公募開始後、参加表明書・提案書の提出、12月7日の総合評価委員会などを経て、最終的に2社と契約書を締結した。

3.4.1 問題の構成

2社それぞれの問題構成の概要を表3にまとめた。両社とも読むこと「聞くこと」「読むこと」「書くこと」に関する試験が100~110分、「話すこと」に関する試験が10分となっている。問題数は、A社は86問、B社は65問であった。

表3 2社の問題構成

	A社		B社	
	問題数	解答時間	問題数	解答時間
聞くこと	30問	25分	20問	20分
読むこと	43問	45分	40問	50分
書くこと	5問	31分	2問	40分
話すこと	8問	10分	3問	10分 (2人1組)
全体	86問	111分	65問	120分

3.4.2 A社の問題イメージの内容

A社の問題イメージについて、技能ごとに問題の概要、問題数、解答形式、CEFRレベルをまとめたものを表4～表7に示す。なお、A社の問題イメージならびに実施や採点方法の解説を含む報告書は別冊に収録してある。

①聞くこと

実施時間：25分

表4 A社の聞くことの問題イメージ

問題の概要	問題数	解答形式	CEFR レベル
①写真・イラスト説明問題：写真やイラストなど与えられた視覚的情報をもとに、ある状況や場面、事物を描写説明した単文レベルの英文を正しく聞き分ける。	7	多肢選択	A1～A2
②会話応答問題：事前予測ができる情報がない中で、会話的な不意の問いかけに対する適当な応答英文を素早く判断し、処理する。	7	多肢選択	A1～A2
③課題解決問題：日本語で事前に与えられる状況設定および視覚的情報と音声情報から、その場面で求められている課題（タスク）を解決する。	8	多肢選択	A2～B1
④要点理解問題：一定以上の長さの英文音声の中から、事前に与えられる英文の質問に答えるために必要な情報を選択して引き出し、求められている解答を導く。	8	多肢選択	A2～B1

②読むこと

実施時間：45分

表5 A社の読むことの問題イメージ

問題の概要	問題数	解答形式	CEFR レベル
①語彙・語法問題：単文レベルの英文の中で、文脈的なつながりを理解し、かつ文法や語彙として最も適切な表現を正確に判断する。	14	多肢選択	A2～B1
②概要把握・情報検索問題 ●概要把握問題：英語の新聞や雑誌の記事を模した題材から短時間で全体の概要を理解する。 ●情報検索問題：英語の広告や案内物などを模した題材から短時間で必要な情報を引き出す（情報検索）。	14	多肢選択	A2～B1
③要点・詳細理解問題：まとまった量の英文に対して、英文の主旨に関する内容や詳細部分の要点について、英文を理解し、必要な情報を読み取る。	15	多肢選択	B1

③書くこと

実施時間：31分

表6 A社の書くことの問題イメージ

問題の概要	問題数	解答形式	CEFR レベル
①質問用紙記入問題：与えられた英文と状況設定を読み、条件にあった内容を解答する力を測定する問題。身近な問いに対して適切な英語で解答する。	3	10語程度の記述	A1～A2
②意見展開問題：社会との接点を通して、個人の経験や他の事例を元に自分の意見と理由を述べる。（事実と意見を区別して書く）	1	100語程度の記述	A2～B1
③情報の要約問題[技能融合問題]：聞いた情報を指定語数で要約する。読み上げられる問題の要点を内容を踏まえて判断し、情報を誤りなく簡潔に表現する。耳から聞いた情報を書いて要約する、リスニングとライティングの技能融合問題。	1	30語程度の記述	B1

④話すこと

実施時間：10分（PCやタブレットによって一斉実施）

表7 A社の話すことの問題イメージ

問題の概要	問題数	解答形式	CEFR レベル
①音読：与えられた英文を正しい発音・リズム・イントネーションで音読し、読み上げた英文の状況や概要を理解する。	2	・80語程度の音読 ・10語程度の口述	A2～B1
②質問を聞いて応答する：図示された情報を読み取り、それに関する質問を聞き取ったうえで、適切に応答する。	4	10語程度の口述	A1～A2
③ストーリーを英語で話す：日常的な出来事について、話の流れを踏まえて相手に伝わるように状況を説明する。	1	40語程度の口述	A2～B1
④自分の意見を述べる：身近で社会的なテーマに対して、自分の意見とその意見の理由を言う。	1	60語程度の口述	A2～B1

3.4.3 B社の問題イメージの内容

B社の問題イメージについて、技能ごとに問題の概要、問題数、解答形式、CEFRレベルをまとめたものを表8～表11に示す。なお、実施や採点方法の解説も含めた、B社の報告書は別冊に収録してある。

①聞くこと

実施時間：20分

表8 B社の聞くことの問題イメージ

問題の概要	問題数	解答形式	CEFR レベル
①短い対話文から重要な情報を取り出し、適切なイラストや写真を選択する。	5	多肢選択	A1～A2
②長めのインタビューを聞いて重要な情報を聞き取る問題。	5	多肢選択	A2～B1
③短い文章を聞き、全体の要点や意味を理解する問題。	5	多肢選択	A2～B1
④長めのモノローグを理解し、空欄に数語の単語を埋める。	5	数語の記述	A2～B1

②読むこと

実施時間：50分

表9 B社の読むことの問題イメージ

問題の概要	問題数	解答形式	CEFR レベル
①非常に短いテキスト（看板、メッセージ、メモ、ラベル等）から伝えたいメッセージを読み取る。	5	多肢選択	A2
②トピックやストーリーライン（話の筋）の結合：文章を読んで適切な語彙を識別する。	5	多肢選択	A2
③実際の新聞記事や雑誌記事を出典とする短いテキストを読み、詳細を理解し主旨を読み取る。	5	多肢選択	A2～B1
④マッチング問題：説明分と合致するものを実際のテキストからとってきた短い文章から選ぶ問題。内容の詳細を理解する必要がある。	5	多肢選択	B1
⑤実際の長いテキストを読み、書き手の伝えたいこと、要点、全体的な意味を理解し、読み取る。	5	多肢選択	B1
⑥実際の長いテキストを読み、単語が文がテキストの中でどのような意味をもつのかを理解し、読み取る。	5	多肢選択	B2
⑦語彙や文法について理解し、適切な語を選択し、短いテキスト中の空欄を埋める。	5	多肢選択	B1
⑧文章の構造などを理解し、空欄を埋める。	5	1語の記述	B1

③書くこと

実施時間：40分

表10 B社の書くことの問題イメージ

問題の概要	問題数	解答形式	CEFR レベル
①短い説明文を読み、説明を踏まえながらeメールを書く。	1	40語程度の記述	A2～B1
②雑誌に掲載される広告を自由に書く。	1	100語程度の記述	A2～B2

④話すこと

実施時間：2人で10分（2人1組で受験し、2人の試験官による面接）

表 11 B 社の話すことの問題イメージ

問題の概要	問題数	解答形式	CEFR レベル
①受験者それぞれが試験官と会話をする。試験官は受験者に交互に質問をする。	1	2人で3分の口述	A2～B1
②試験官が絵や写真などを与えながら状況を説明した後、受験者同士で話し合う。	1	2人で3分の口述	A2～B2
③試験官が②のトピックに関連した質問をし、受験者2名とディスカッションする。	1	2人で4分の口述	A2～B2

3.5 まとめと今後の課題

英語については、公募によって民間事業者等の知見を得ながら、問題イメージ等を作成することになった。その結果、4技能のうち、特に現在のセンター試験で実施されていない「書くこと」と「話すこと」に関しては、2社の問題イメージから様々な知見が得ることができた。

3.5.1 問題の題材と難易度

2社ともに問題の題材としては、アカデミックな内容より、高校生にとって身近な内容が中心であった。また、細かく設定された条件にしたがって解答する問題と、少ない条件のもと自由に解答する問題とがあり、後者の方が難易度は高かった。難易度に関して、文部科学省が高校3年生を対象に実施している英語力調査において、「読むこと」「聞くこと」ではCEFRのA1上位からA2下位レベルに集中していたものの、「書くこと」「話すこと」ではA1下位レベルが多いと報告されている。そのため、特に「書くこと」「話すこと」に関しては、どの学力層をコアターゲットとするかを検討するとともに、実証的な検証も必要である。

3.5.2 話すことの実施方法

「話すこと」の実施方法に関して、A社はPCなどによる一斉実施なのに対して、B社は2人1組で受験し、2人の試験官が面接を行うものであった。2社とも、それぞれの方法に関して十分な実績がある。しかし、同時または短期間での50万人の「話すこと」に関する試験は世界でも前例がない。2社の提案を参考にしつつ、大規模で行う場合にはどのような方法なら実現可能か、実施ならびに公平な採点が限られた費用や時間内で可能かなどの点を、モニター調査などによって検証することが必要である。

4. 試作問題評価のためのモニター調査の実施

4.1 モニター調査の実施目的と概要

このモニター調査の目的は、現在議論が進められている「大学入学希望者学力評価テスト(仮称)」に向けて新しく作成した試作問題を、現行のセンター試験を受験している大学1年生に出題し、解答させることによって、難易度に関する情報や、記述式問題に対する実際の解答例を収集することである。そして、それらの情報を作題部会にフィードバックし、内容に関する議論を深め、記述式問題の採点基準の策定などに資することをめざしている。

モニター調査は、実施時期別に大きく3回に分けて実施された。

第1回 モニター調査：

- ① 平成27年10月31日(土): SA大学(293名)
- ② 平成27年11月14日(土): CH大学(192名)

第2回 モニター調査：平成28年1月17日(日), 24(日): NTT DATA 駒場研修所(363名)

参加大学: A大学・B大学・C大学・D大学・E大学

(大学入試センター研究開発部による本追モニター調査内の実験調査枠にて実施)

第3回 モニター調査：平成28年2月17日(水): NTT DATA 駒場研修所(64名)

参加大学: A大学・B大学・C大学・D大学・E大学

(第2回の調査参加者の中から参加希望者を募集して実施)

ここで対象とする大学の選択に至った背景について述べる。第1回のモニター調査では、SA大学やCH大学といった、地方の国立総合大学の1年生を対象としている。当初、高校の3年生を対象にする案も出されたところであるが、現在の高校生は現行のシステムで大学を受験することになっている。したがって、未知の新しい傾向の試験を不用意に持ち込んで、高校現場に不要な混乱を生じせしめるのは避けなければならない。そこで検討の結果、高校で彼らと同じ新教育課程で学んできた大学の1年生を対象とすることとした。

また彼らは、主にセンター試験の成績で選抜されている層であり、センター試験の中核的な受験者層に相当する。さらに、両大学は総合大学なので、教員養成系から理学部や医学部まで設置されており、より広範囲の学力層を対象にすることができると考えた。

それに対して、第2回、第3回のモニター調査は、東京都内の選抜性が極めて高い大学の学生を対象とした。彼らは、本事業のスタートの時点で提起されていた、より高度な内容の試験問題の検証も可能と考えられる受験者層であるため、調査計画に盛り込んで実施した。

4.2 モニター調査での試験の内容

試験の問題内容の主な概略は次の通りである。

第1回と第2回のモニター調査では、2章で詳細な説明のあった「新テスト試作問題[国語、数学、地理歴史(世界史)、理科(物理)]」の試験を行った。また、CBTシステムの実施検証、および記述式問題の採点に係わる資料収集を目的とし、CBTによる評価試験と記述式による、一般教養・複合領域の試験も行った。

第2回のモニター調査は、研究開発部の本追モニター調査の一環として行われているので、参加者は平成28年度センター試験の全ての教科についても受験している。ここではさらに、平成27年12月に公表された「大学入学希望者学力評価テスト(仮称)」の数学に関する問題イメージ例の中の、「スーパームーン問題」についても合わせて実施した。

最後の第3回のモニター調査では、現在、思考力や判断力を測っているとされている既存の試験問題を利用して実施した。具体的には下記のような問題である。

- ① 久原・井上・波多野(1983)による「批判的思考力テスト」
- ② PISAの過去問題(読解力・数学的リテラシー・科学的リテラシー)
- ③ 法科大学院適性試験の「論理的判断力」、「分析的判断力」、「長文読解力」の問題
- ④ 教育政策研究所による特定課題研究の中の「論理的思考」の問題

さらに、国語の能力によって問題文の読み方に違いがあるかを明らかにするため、国語の問題イメージ例を使用して、アイ・トラッカーを用いた注視点の追跡実験も並行して行った。

この第3回のモニター調査の参加者は、第2回の調査に参加した者の中から募集している。したがって、上記の思考力や判断力を測る試験問題と、第2回の調査の時に受験した現行のセンター試験の成績との関係や、さらに新しく作成された試作問題の成績との関係についても検討することが可能となる。

4.3 第1回モニター調査の実施状況

4.3.1 SA大学におけるモニター調査

調査実施日時 平成27年10月31日(土) 参加者数: 293名

参加応募期間 10月8日(木)~10月23日(金) 応募総数: 326名

募集の周知方法 全学の1年生を対象とした英語の授業にて全員に募集案内を配布した(10/8-19)。なお募集開始時には、新テストの方向性がまだ固まっていなかったため、不要な物議を醸すことがないように、調査目的をぼかし「大学生の学力に関する調査」とした(図1)。

応募方法 募集案内に記載されたQRコードをスマートフォンなどで読取って、web上に準備した申込みページにアクセスする。入力フォームに必要事項を記載して参加の申込みを行う(図2)。その後、メールによって、受験案内を送付した(図3, 図4)。

試験の時間割と科目構成 時間割と科目を表1に示す。この調査の構成で特徴的なのは、「一般教養・複合領域」、「国語領域」、「数学領域」、「理科領域(物理)」の科目において、参加者を2分割し、異なる形式の試験に解答させているところである。たとえば「一般教養・複合領域」では、参加者の半数はタブレットPCを用いたCBTによる試験を受け、別教室のもう半数の者は、紙筆方式の記述問題に解答している。

表 1 第 1 回モニター調査の時間割と科目の構成

時間		科目	
10:30-12:00	90分	一般教養・複合領域 【CBTによる試験】 【筆記式記述問題】	
12:00-13:00	60分	<昼休み>	
13:00-14:30	90分	国語領域 【J1: 題材Aの問題】 【J2: 題材Bの問題】	
14:30-14:40	10分	<休憩>	
14:40-16:10	90分	数学領域 【M1: 選択式問題】 【M2: 筆記式問題】	
16:10-16:20	10分	<休憩>	
16:20-17:50	90分	地歴領域(世界史)	理科領域(物理) 【P1】 【P2】

SA 大学の参加者の割当て 国語や数学等で異なる問題冊子が使用されるので、それぞれの冊子を受験する参加者の学部構成比率に偏りが出ないように、学部別に層化して参加者を割り当てた(表 2)。また、世界史は高校で必修科目になっているので、センター試験で世界史を受験していないものの受験も認めている。しかし、現実的には実際に受験の準備をした者と、しなかった者とは学力に開きがあり、試験に向かう姿勢も異なることが予想される。そこで、両者が同じ試験室に混在しないように、異なる試験室に割り当てた。

SA 大学の参加者 調査の当日、実際に参加したのは計 293 名であった。参加率で見ると、89.9%であり、ほぼ 9 割の参加状況であった。学部別に参加者をまとめたものを表 3 に示す。

表 2 SA 大学における参加者の割当て

試験室	人数	1時限: 一般教養		2時限:国語 3時限:数学		4時限: 世界史		4時限: 物理				学部				
		CBT方式	紙筆方式	J1/M1 冊子	J2/M2 冊子	セ試 受験	未 受験	冊子 P1	冊子 P2	新課程	旧課程	教育学部	教養学部	経済学部	工学部	理学部
301	53	0	53	53	0	0	53	0	0	0	0	13	6	19	1	14
302	60	0	60	60	0	0	0	0	60	44	16	5	0	1	50	4
401	54	0	54	0	54	0	54	0	0	0	0	13	5	20	2	14
402	59	0	59	0	59	0	0	59	0	43	16	5	0	1	50	3
501	52	52	0	26	26	52	0	0	0	0	0	12	10	27	3	0
502	48	48	0	24	24	0	0	24	24	48	0	0	0	0	21	27
合計	326	100	226	163	163	52	107	83	84	135	32	48	21	68	127	62

表 3. SA 大学における学部別の調査参加者の人数と参加率

教育学部	42 名	87.5%
教養学部	21 名	100.0%
経済学部	60 名	88.2%
工学部	115 名	90.6%
理学部	55 名	88.7%
合計	293 名	89.9%

4.3.2 CH 大学におけるモニター調査

調査実施日時 平成 27 年 11 月 14 日(土) 参加者数: 192 名

参加応募期間 10 月 13 日(火)～11 月 9 日(月) 応募総数: 186 名

募集の周知方法 学務部長が各学部の学部長に主旨を説明し、定員に比例した人数が参加してもらえよう個別に依頼する形で募集した。

応募方法 応募希望者には SA 大学と同様の募集案内を渡し、web の入力フォームに必要な事項を記載して参加申込みを行った。その後、手順は SA 大学と同様である。

試験の時間割と科目構成 時間割と科目についても、先の表 1 の通りである。

CH 大学の参加者の割当て 各科目の冊子の割当て条件に則して、学部別に層化して参加者を割り当てた。

CH 大学の参加者 参加者は 192 名であった。これは、事前の応募総数よりも多い人数となっている。これは、周知方法が個別に依頼する形であったため、事前の応募がない状態で、直接試験会場に来た参加者が現れたためである。これらについては、予備の冊子を利用し、追加で受験番号を発行して対応した。最終的な学部別の参加者数を表 4 に示す。

表 4. CH 大学における調査参加者の学部別の人数

文学部	教育学部	法政経学部	工学部	理学部	看護学部	薬学部	医学部	園芸学部	合計
10 名	33 名	42 名	51 名	26 名	3 名	11 名	7 名	9 名	192 名

4.4 第 2 回モニター調査の実施状況

第 2 回のモニター調査は、平成 28 年 1 月 16 日(土)・17 日(日)、23 日(土)・24 日(日)に、大学入試センターに隣接する NTT DATA 駒場研修所で行われた。この第 2 回調査は、研究開発部が継続的に実施している本追モニター調査の中の実験調査時間枠の中で行われた。なお本追モニター調査の詳細は、大学入試センター研究開発部(2016)の報告を参照されたい。

ここではその中の、第 2 回モニター調査に関連する部分のみについて報告する。

調査実施日時 平成 28 年 1 月 17 日(日)、24 日(日): 参加者数: 363 名

参加応募期間 平成 27 年 10 月～12 月 応募総数: 396 名

募集の周知方法 東京都内の 5 つの国立大学(A 大学・B 大学・C 大学・D 大学・E 大学)の 1 年生を対象とした。各大学でのアルバイトを斡旋する窓口を通して参加者を募集した。

応募方法 応募希望者は所定の書類を作成し、各大学の窓口提出して申し込んだ。

試験の時間割と科目構成 本モニターに関係する箇所々の時間割について表5に示す。

参加者の割当て 各科目ごとの割当て条件に則して、まず大学別に層化した上で、さらに文系理系に区分してさらに層化して上で、参加者を試験室に割り当てた(表6)。

大学別の参加者 最終的に4日間の本追モニター調査に全日出席した最終的な参加者は363名であった。その参加者の大学別の参加人数と参加率を表7に示す。

第3回モニター調査の募集 モニター調査の最終日の全日程を終了した後に、この調査の参加者のみを対象とし、第3回モニター調査参加への募集案内を行った。

表5 第2回モニター調査の時間割と科目の構成

実施日	時間		科目		
1月17日(日)	12:35-14:00	85分	実験A 地歴領域 (世界史)	実験B* 複合領域C (CBT)	実験C* 複合領域P (紙筆試験)
	14:25-15:50	85分	実験D 数学領域 【M3冊子: 選択式問題】 【M4冊子: 筆記式問題】		
1月24日(日)	12:35-14:00	85分	実験B* 複合領域C (CBT)	実験F 理科領域 (物理 [P3/P4冊子])	
	14:25-15:50	85分	実験G 国語領域 【J3冊子】		

* 実験 B,C の中で数学の問題イメージ例の中の「スーパームーン問題」も実施

表6 第2回モニター調査の実験科目の割り付け結果

	実験 A 世界史	実験 B CBT 1/17	実験 B' CBT 1/24	実験 C 紙筆式	実験 D 数学	実験 F 物理	実験 G 国語
A 大学・文系	46	1	35	10	57		57
B 大学 (文系)	30	10	28	21	61		61
D 大学・文系	13	11	10	11	35		35
文系小計	89	22	73	42	153		153
A 大学・理系	7	17	5	74	98	61	98
C 大学 (理系)	10	5	8	52	67	41	67
D 大学・理系	6	8	3	22	36	14	36
E 大学 (理系)	8	4	3	30	42	28	42
理系小計	31	34	19	178	243	144	243
合計	120	56	92	220	396	144	396

表7 第2回モニター調査の大学別(文理別)の参加人数と参加率

	A 大学 (文系)	A 大学 (理系)	B 大学 (文系)	C 大学 (理系)	D 大学 (教育系)	E 大学 (理系)	合計
参加人数 (参加率)	54 名 (95%)	92 名 (94%)	52 名 (85%)	60 名 (90%)	65 名 (92%)	40 名 (95%)	363 名 (92%)

4.5 第3回モニター調査の実施状況

第3回の調査は、第2回と同じNTT DATA 駒場研修センターで行われた。参加者は第2回の調査の全日程に参加した者の内、この第3回の調査に応募した者である。

調査実施日時 平成28年2月17日(水): 参加者数: 64名

参加応募期間 平成28年1月24日(日)~2月8日(月) 応募総数: 75名

募集の周知方法 第2回調査に参加した5つの国立大学(A大学・B大学・C大学・D大学・E大学)の1年生を対象とし、第2回調査の終了時に募集案内を配付した。

応募方法 募集案内に記載されたQRコード、もしくはurlから申込み用のwebフォームに進み、必要事項を入力して応募した。

試験の時間割と科目構成 第3回調査は、思考力や判断力を測定しているとされる既存の試験問題を用いた。時間割と科目構成について、表8に示す。

参加者の割当て 原則として、参加者は同一の試験問題を受験した(表8)。ただし、一部の参加者については、センター試験の国語の問題、および新テストの問題イメージ例を用いた、アイ・トラッカーによる注視点の追跡実験に参加した。

大学別の参加者 調査に出席した参加者は64名であった。大学別の参加人数を表9に示す。

表8 第3回モニター調査の時間割と科目の構成

時間	試験科目 (試験時間)
10:20-10:40	【批判的思考力テスト】 (20分)
10:40-12:00	【総合領域 A: PISAの過去問題】 (75分)
≪ 昼休み ≫	
13:00-13:45	【総合領域 B*: 論理的判断力】 (40分)
[休憩]	
13:55-14:40	【総合領域 C*: 分析的判断力】 (40分)
[休憩]	
14:50-15:35	【総合領域 D*: 長文読解力】 (40分)
[休憩]	
15:45-17:30	【総合領域 E: 論理的な思考 (国研:特定課題)】 (100分)

* 一部の参加者(6名)はアイ・トラッカーを用いた注視点の追跡実験に参加

表 9 第 3 回モニター調査の大学別の参加人数と参加率

	A 大学 (文・理系)	B 大学 (文系)	C 大学 (理系)	D 大学 (教育系)	E 大学 (理系)	合計
参加人数 (参加率)	38 名 (86%)	11 名 (100%)	10 名 (83%)	4 名 (50%)	1 名 (100%)	64 名 (85%)

参考文献

大学入試センター研究開発部(編)(2016). 平成 27 年度 大学入試センター試験 モニター調査
研究報告書.

5 モニター調査の分析結果

5.1 国語

5.1.1 問題冊子の構成（第1回、第2回）

第1回モニター調査で用いたのは、J1冊子とJ2冊子の2種類である。表1と表2に、それぞれ、J1冊子とJ2冊子の問題構成を示す。いずれも、解答時間は80分である。J1冊子とJ2冊子は、別冊に収録する。

表 1: 国語 J1 冊子の構成

大問	題材	設問	解答形式
第1問	3つの文章 (イメージ⑦)	問1	X: 20字以内の記述 Y: 20字以内の記述
		問2	2つの文の誤り箇所を正しく直す
		問3	連動型複数選択式（状況、問題、解決） 多肢選択肢の組合せ
第2問	複数の国語辞典 (イメージ④)	問1	選択肢の組合せ
		問2	多肢選択式（2つ選ぶ）
		問3	(1) 多肢選択式 (2) 15字以内の記述
第3問	古典に関する近代以降 の文章(イメージ③)	問1～2, 4～7	多肢選択式
		問3, 8	多肢選択式（2つ選ぶ）

表 2: 国語 J2 冊子の構成

大問	題材	設問	解答形式
第1問	3つの文章	問1～2	多肢選択式
		問3	連動型複数選択式（状況、問題、解決） 選択肢の組合せ
第2問	交通事故に関するグラフと 生徒の話し合い（会話）	問1	ア、イ 各 15字以内の記述
		問2	(1) 多肢選択式 (2) 記述（文字数指定なし） (3) 25字以上 30字以内の記述
		問3	記述
		問4	20字以内の記述
第3問	平成24年度センター試験 「国語」第4問(本試験) をもとに変更(イメージ①)	問1	ア～ウ：多肢選択式（要約文の完成）
		問2	エ～カ：多肢選択式（内容整理図） キ：40字以内の記述 キの根拠：多肢選択式

J1冊子の第1問とJ2冊子の第1問は、題材は異なるが、いずれも、3つの文章を比べて読ませ

ることで、思考力・判断力・表現力を測定しようという意図で作成されている。それぞれ、問3が連動型複数選択式の設問である。

J1冊子の第2問では、複数の国語辞典の同じ語の項目を資料として示される。それぞれの表現の特徴や、内容・構成の共通点や相違点をとらえたり、それをもとに、根拠を明確にして意見を文章にまとめたりする能力などを問う。J2冊子の第2問では、交通事故の発生件数、負傷者数、死者数の推移を示したグラフと高校生の話し合いの記録が示される。これらをもとに、事実関係の把握や推測、仮説を検証するために必要な資料の選択、その資料を基にした推論や根拠などについての考察ができるかを問うと共に、事象の関係性について洞察して予測をたて、自分の考えを適切に説明することができるかを問う。

J1冊子の第3問は、古典に関する近代以降の文章という新傾向の素材文を用いて、思考力・判断力・表現力を問う。J2冊子の第3問は、平成24年度センター試験（本試験）の第4問（漢文）をもとに作成されている。問題文全体を要約したり、話の意図を根拠に基づいて推論したりする能力を問う。

第1回モニター調査では、部会で作成された問題をなるべく多く用いて解答データを得る目的で、受験者を同質とみなせる2つの集団に分割して、J1冊子とJ2冊子を割り当てた。解答終了後のアンケートでは、解答方法の説明のわかりやすさ、解答時間が十分だったか、解答用紙のスペースが十分だったか、各大問の難易度に関する評定や、難しかった点や題意がわかりにくかった点に関する自由記述による回答を収集した。

第2回モニター調査で用いたのは、J3冊子である。表3に、J3冊子の問題構成を示す。解答時間は80分である。J3冊子は、別冊に収録する。いずれの大問も、平成27年12月22日の高大接続システム改革会議において、新テストの記述式問題イメージ例（国語）として公表されたものである。ただし、著作権の関係で、題材の一部は公表されていない。

表 3: 国語 J3 冊子の構成

大問	題材	設問	解答形式
第1問	交通事故に関するグラフと生徒の話し合い（会話）	問1	40字以内の記述
		問2	80字以上100字以内の記述
第2問	3つの文章	状況	選択式
		問題	選択式
		解決	30字以上50字以内の記述
第3問	新聞の社説		200字以上300字以内の記述

第1問は、「多様な見方や考え方が可能な題材に関する複数の図表や文章を読み、情報を統合しながら、考えを構成し表現する問題」（高大接続システム改革会議, 2015）の例として示された問題である。問1, 問2共に、話し合いにおける発言の一部を記述することが求められる。

第2問は、「3つの文章で語られている状況、問題、解決法に関する共通点について考察し、選択式と記述式で構成・表現する問題」（高大接続システム改革会議, 2015）の例として示された問題である。3つの文章で述べられた「状況」「問題」「解決」法の共通点について考察することが求められる問題であり、「状況」と「問題」についてはそれぞれの選択肢群から1つを選び、それ

を踏まえた「解決」を記述することが求められる。「状況」と「問題」の選択肢の組合せは1通りではなく、それらの組合せに応じて複数の「解決」法が成立するという特徴を持つ。

第3問は、「1,400字程度の新聞記事を、一定の目的に添って読み取り、得られた情報を取捨選択したり、自分の考えを統合したりしながら、新たな考えにまとめ、200～300字で表現する問題」(高大接続システム改革会議, 2015)の例として示された問題である。

表4に、第1回モニター調査で用いた問題と第2回モニター調査で用いた問題の対応、および、変更点を示す。J3冊子の3つの大問のうち、第1問と第2問については、第1回モニター調査で試行した問題をもとに、記述させる文字数を長くしたり、解答形式を選択式から記述式に変更したり、といった変更が加えられている。例えば、J3冊子第1問の間1では「空欄アに当てはまる内容を40字以内で」記述するよう指示されるが、この箇所は、J2冊子第2問の素材文の空欄イ(問1で15字以内で記述するよう指示された箇所)にその直後の部分を加えた箇所に該当する。第3問は、新規の問題である。

表 4: J3 冊子の問題と第 1 回モニターの問題との対応

J3 冊子の問題		第 1 回モニター調査の問題との対応		変更点
第 1 問	問 1	J2 冊子第 2 問	問 1	「空欄イ」の後ろまで範囲を広げた箇所について記述させる
	問 2		問 2	「空欄エ」「空欄オ」とその周辺部分まで範囲を広げて記述させる
第 2 問	「解決」	J1 冊子第 1 問	問 3	「解決」を多肢選択式から記述式に変更
第 3 問		新規		

第2回モニター調査では、なるべく多人数の解答データを収集するのを目的として、受験者全員にJ3冊子を解答させた。解答終了後のアンケートでは、解答方法の説明のわかりやすさ、解答時間が十分だったか、解答用紙のスペースが十分だったか、各大問の難易度に関する評定や、難しかった点や題意がわかりにくかった点に関する自由記述による回答を収集した。

5.1.2 第1回モニター調査に関する分析

多肢選択式の設問については、受験者が解答用紙に記入した選択肢番号を入力業者が電子データ化したものを採点した。記述式の設問については、作題グループの委員による採点が行われた。各大問を作成した委員が、解答類型を作成して、採点を行った。

5.1.2.1 連動型複数選択式の設問

連動型複数選択式の設問であるJ1冊子第1問の間3では、「該当する組合せは一つとは限らないが、あなたが答えるのは一つでよい」と指示されている。正答とされる組合せ2つのうち、「2-4-2」と解答した者の割合は29.3%、「2-6-4」と解答した者の割合は17.8%であり、両方を合わせると47.1%が正答した。

同じ連動型複数選択式の設問であるJ2冊子第1問の間3では、「該当する組合せは二つとは限ら

ないが、あなたが答えるのは二つでよい」と指示されている。正答とされる組合せは「1-2-3」「1-6-6」「3-6-6」の3つである。解答した2つの組合せの両方正答の者の割合は7.0%であった。どちらかの組合せが正答の者の割合は26.7%であった。解答の順番は問わずに、1つめまたは2つめの組合せとして「1-2-3」と解答した者の割合は18.5%、「1-6-6」と解答した者の割合は4.5%、「3-6-6」と解答した者の割合は17.7%であり、仮にこれらを加算すると、正答とされる3つの組合せのうちどれか1つを解答できた者の割合は、40.7%となる。

J1 冊子第1問と J2 冊子第1問では、異なる文章が題材として用いられているので、各冊子の連動型複数選択式の設問の正答率を直接比較することは出来ないが、正答とされる組合せのどれか1つを解答できた者の割合という点でみると、J1 冊子第1問の間3も J2 冊子第1問の間3も、受験者の約4~5割程度である。

各冊子の連動型複数選択式の設問の難易度について、アンケートで「やや難しかった」または「難しかった」と回答した者の割合は、J1 冊子では33.1%、J2 冊子では66.3%である。自由記述では、連動型複数選択式の設問への解答のしかたのわかりにくさに関する指摘がみられる。それぞれ、「一つとは限らないが」や「二つとは限らないが」という点に関する迷いに加えて、J2 冊子第1問の間3に関しては、「該当する組合せは二つとは限らないが」「答えるのは二つでよい」という点のわかりにくさを指摘する記述が目立つ。「二つ」答える際に同じ選択肢番号を含んでもよいかどうかや、「二つ」答える際の解答順について、問題文に説明がないことに起因する答えにくさが指摘されている。

5.1.2.2 記述式の設問

J1 冊子第2問の間3は、(1)は多肢選択式で、(2)では、その理由を述べた文章の空欄を15字以内で空欄を補って記述する問題である。(1)については受験者の78.9%が正答したが、(2)では、◎は受験者の14.9%、○は受験者の2.5%にとどまる。

J2 冊子第2問には、記述式の設問が多く含まれるが、正答の者の割合は設問によって大きく異なる。J2 冊子第2問の記述式の設問の中で、問2(2)（空欄エ）が最も◎の者の割合が高いが(92.2%)、その直後の問2(3)（空欄オ）では◎の者の割合が最も低い(16.9%)。

J2 冊子第3問では、問2において話の内容を整理した図の空欄エ～オを多肢選択式で解答させると共に、その図に即して登場人物の意図を解釈した文章の空欄キを記述させる。多肢選択式の空欄エ～オについては、正答の者の割合は77.8%~93.4%であったが、記述式の空欄キについては、○の者の割合は10.7%にとどまる。

さらに、空欄キの解答の根拠について、7つの選択肢から「該当するものをすべて」選ぶ設問では、該当する選択肢4つをすべて選択できた者の割合は2.9%にとどまるが、4つのうちいずれか1つが欠けた者（準正答）の割合は18.5%である。いずれか2つが欠けた者の割合は22.2%で、正答と準正答の割合を上回る。正答選択肢がいくつあるか明示されていない状態ですべて選ぶのは、受験生にとって非常に難しいことが示唆される。

5.1.2.3 大問得点間の相関係数

J1 冊子と J2 冊子について、以下の仮配点によって大問得点を算出した。表5に、各冊子の大問得点間の相関係数を示す。

- 多肢選択式の設問:正答を1, 非正答を0。組合せ式の設問についてはすべて合っているものを1, それ以外を0。
- 記述式の設問 : 正答 (○) を1, 準正答 (△) を0.5 など。

J1 冊子の大問得点間の相関係数で有意なのは, 第1問と第2問の相関係数0.16と, 第2問と第3問の相関係数0.18である。第1問と第3問の大問得点には有意な相関は認められない。J2 冊子の大問得点間の相関係数で有意なのは, 第1問と第3問の相関係数0.18である。第1問と第2問の大問得点, 第1問と第3問の大問得点には有意な相関は認められない。このように大問得点の相関が非常に低い状態の場合, これらの大問得点を合計して冊子得点を算出しても, その値を一定の方向性を持つ能力の指標として用いるのは適切ではないことが示唆される。

表 5: 国語の大問得点間の相関係数 (*: $p < 0.05$, **: $p < 0.01$)

(a) J1 冊子

	第1問	第2問	第3問
第1問	1	0.16 *	0.03
第2問	0.16 *	1	0.18 **
第3問	0.03	0.18 **	1

(b) J2 冊子

	第1問	第2問	第3問
第1問	1	0.10	0.18 **
第2問	0.10	1	0.02
第3問	0.18 **	0.02	1

5.1.3 第2回モニター調査に関する分析

多肢選択式の設問については, 受験者が解答用紙に記入した選択肢番号を入力業者が電子データ化したものを採点した。記述式の設問については, 作題グループの委員が分担して採点を行った。

第2回モニター調査で用いた J3 冊子の問題と, センター試験の得点や思考力を測るテストとの相関については, 5.5 節で述べる。ここでは, 各設問についての分析結果を述べる。

5.1.3.1 第1問

J3 冊子第1問は, 問1が40字以内, 問2が80字以上100字以内の記述である。問1で○の者の割合は41.8%であるのに対して, 問2で◎の者の割合は0.5%, ○の者の割合は1.6%にとどまる。

J3 冊子第1問は, 第1回モニター調査の J2 冊子第2問に対応する。J3 冊子第1問の空欄の箇所や解答字数は, J2 冊子第2問 (第1回モニター調査) と少し異なるが, 対応する設問の正答率を比べてみる。

J3 冊子第1問の問1に対応する J2 冊子第2問の問1空欄イ (15字以内の記述) では, ◎の者の割合は81.1%である。J3 冊子第1問の問2に対応する J2 冊子第2問の問2空欄エ (文字数指定なし), 空欄オ (25字以上30字以内の記述) では, ◎の者の割合は, それぞれ, 92.2%, 16.9%で

ある。すなわち、J3 冊子第1問では、問1も問2も、正答の者の割合が非常に低くなっている。

表4に示したように、J3 冊子第1問の各設問の空欄は、J2 冊子第2問で対応する空欄を広げた範囲に該当する。記述させる範囲が広がったことで、受験者の解答が、J2 冊子で正答とした範囲よりも多様な内容を含むようになったとも考えられる。

5.1.3.2 第2問

J3 冊子第2問は、連動複数選択式の設問で、「状況」「問題」は多肢選択式、「解決」は30字以上50字以内の記述である。正答とされる「状況2-問題4」を選んだ者の割合は受験者の40.7%で、そのうち「解決」の採点結果が○の者は53.4%であり、全受験者に占める割合は21.7%である。正答とされる「状況2-問題6」を選んだ者の割合は受験者の16.5%で、そのうち「解決」の採点結果が○の者は48.3%であり、全受験者に占める割合は8.0%である。

J3 冊子第2問は、第1回モニター調査のJ1 冊子第1問の間3に対応する。J1 冊子第1問の間3とJ3 冊子第2問では、「状況」「問題」の選択肢はほぼ同じであり、「状況-問題」の組合せを正しく選べた者の割合は、どちらも同程度（60%前後）であるが、そのうち「解決」でも正答を得た者の割合は異なる。J1 冊子第1問の間3（第1回モニター調査）では、「状況」と「問題」の組み合わせを正しく選んだ者のうち70~80%が「解決」の選択肢を正しく選んでいる。それに対して、J3 冊子第2問（第2回モニター調査）では、「状況」と「問題」の組み合わせを正しく選んだ者のうち、「解決」の記述で○と採点された者の割合は50%程度である。

連動型複数選択式の大問（第2問）については、J3 冊子においても、解答のしかたのわかりにくさが指摘されている。

5.1.3.3 第3問

J3 冊子第3問は、第2回モニター調査で新規に出題した200~300字の記述式の問題である。「内容」「表現」「構成・論理」という3つの観点で採点が行われた。「内容」のランクが3または4の答案についてのみ、「表現」と「構成・論理」の観点の採点が行われた。「内容」がランク4の者の割合は40.9%、ランク3の者の割合は11.3%である。「表現」ランクと「論理・構成」ランクを間隔尺度とみなして算出した相関係数は0.63 ($p < 0.0001$) である。

アンケートの自由記述で最も多く指摘があった大問は第3問である。特に、「二段落構成」（条件2）で解答することと、「二百字以上、三百字以内」（条件1）で解答することに関して、段落分けした場合の字数の数え方が問題文で明示されていない点が多く指摘されている。さらに、段落の冒頭の字下げ分を文字数に含めるのかどうかについても、問題文で明示されていない点が指摘されている。また、「本文中から引用した言葉」（条件3）については、どの範囲を「引用」とするのかがわかりにくいという指摘がみられる。

5.1.4 まとめ

「状況」「問題」「解決」の選択肢の組合せを解答させた連動型複数選択式の設問では、正答とされる組合せのどれか1つを解答できた受験者は約4~5割程度だが、組合せを2つ解答させたJ2 冊子では、どちらも正答だった受験者は7%にとどまる。アンケートでは、正答とされる組合せが「二つとは限らないが」「答えるのは二つでよい」という指示への戸惑いも指摘されており、新しい解答形式の設問をはじめて解く受験者の心理的負担がうかがえる。第2回モニター調査で

は、「解決」を30字以上50字以内の記述式に変更して出題されたが、「状況」「問題」の組合せを正しく選べた者のうち、「解決」も正しく記述できた者は半数程度である。「解決」を多肢選択式で出題した第1回モニター調査では、「状況」「問題」の組合せを正しく選べた者のうち7～8割が「解決」でも正答を得ていることから、記述式に変更したことで難度が上がったことが示唆される。

記述式の設問については、第1回と第2回モニターを通じて、正答率が1割を切る設問もあれば正答率が9割を超える設問もあり、難度は多様であった。第2回モニター調査において、記述させる範囲を第1回モニター調査よりも広げた設問では、範囲の拡大に応じた採点基準の再検討が必要と思われる設問もみられる。また、第1回と第2回モニター調査の採点は、作題グループの委員によって行われたが、作題者でない者が作題意図に沿った採点をするための採点基準については、今後の検討課題である。

また、多肢選択式でも、正答の個数を知らせずに「該当するものをすべて」選ぶよう指示した設問では、過不足なくすべて選べた受験者の割合は数%程度にとどまった。受験者の理解度に応じた評価の在り方を検討する必要があるだろう。

参考文献

高大接続システム改革会議 (2015). 第9回配付資料 別紙3『「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】』(平成27年12月22日), < http://www.mext.go.jp/b_menu/shingi/chousa/shougai/033/shiryo/_icsFiles/afielddfile/2015/12/22/1365554_06_1.pdf > (2016年2月22日確認).

5.2 数学

5.2.1 問題の冊子構成（第1回、第2回）

数学部会では、「大学入学希望者学力評価テスト（仮称）」（以下、新テスト）でどのような問題が出題可能であるかを検討するための情報を収集することを目的として、2回のモニター調査のいずれにおいても記述式とマーク式の解答形式の試験（調査）を実施することとし、計4種類の問題冊子（数学領域M1～M4）が作成された。これらは、従来型の数学的知識を問う問題に加え、できるだけ誘導をなくしたり、数学的知識を現実の世界における問題に応用・活用したりする問題などから構成される。

4種類の問題冊子（以下、冊子）M1～M4は、いずれも解答時間を80分とし、各冊子には4つの大問を含め、各大問の平均解答時間は20分を想定した。また、各大問の満点は25点、各冊子の満点は100点とした。部会で作成した数学の問題の概要を表1に示す。表中の◎印は記述式問題を、○印はマーク式問題を、それぞれ表す。M1～M4冊子は、別冊に収録する。

表 1: 数学の問題の概要

大問 〔中間〕		問題の概要	問題冊子			
			M1	M2	M3	M4
第1問	〔1〕	証明の間違い探しとその修正	◎		◎	
		証明の場合分け		○		○
	〔2〕	2次関数	◎	○		
		論理			◎	○
第2問		人工衛星から見た地平線までの距離	◎	○	◎	○
第3問		点列生成のアルゴリズム	◎	○		
		数値計画法			◎	○
第4問		カーシェアリングの問題	◎	○	◎	○

第1回モニター調査（以下、第1回調査）では、数学領域M1、数学領域M2の2種類の冊子を用いて調査が行われた。M1冊子とM2冊子は、表2に示す4つの大問で構成されている。現在実施されているセンター試験と比較して、思考力、判断力、表現力の測定を強化することを目標として作成されている。

また、数学的知識を問うことに加え、現実の世界における問題に数学的知識を応用・活用する出題もなされている。新傾向の問題として、証明問題における解答の誤りを発見してその誤りを修正する問題（M1冊子第1問〔1〕）、与えられたアルゴリズムを理解して、設問に解答する能力を見る問題（M1・M2冊子第3問）、「現実の世界」（カーシェアリング）での問題解決を題材とした確率の理解とデータの解釈（M1・M2冊子第4問）などの問題が出題されている。

解答形式は、M1冊子とM2冊子で異なる。今回のモニター調査の目的のひとつは、同一または同様な内容の問題について、記述式とマーク式（現在実施されているセンター試験の数学②方式）という異なる解答形式で解答させた場合に、受験者の解答状況の違いを比較・検討することである。M1冊子は、空欄に数式、数値、短文等を記述させる解答形式であるのに対して、M2冊

表 2: 数学の問題冊子 (M1 冊子, M2 冊子) の構成

大問 〔中間〕		問題の内容	解答形式	
			M1	M2
第 1 問	〔1〕	(M1) 証明の間違い探しとその修正 (M2) 証明の場合分け	数式や短文の 記述	数学②方式 (マーク式)
	〔2〕	2 次関数 (H24 センター試験 数学 I・A 第 2 問 (1) まで (の改題))		
第 2 問		「現実の世界」の問題 (人工衛星から見た地平線までの距離) への数学的知識 (円の接線, 極限) の応用・活用	数式や数値の 記述	数学②方式 (マーク式)
第 3 問		<ul style="list-style-type: none"> ● 数学用語の定義・概念の理解 ● 与えられたアルゴリズムの理解とその応用 ● 求積法の理解 (場合分けにより放物線と直線で囲まれる部分の面積を求める) 	数値の記述	数学②方式 (マーク式)
第 4 問		「現実の世界」(カーシェアリング) での問題 解決を題材とした確率の理解とデータの解釈	数学②方式 (記述式)	数学②方式 (マーク式)

子は、センター試験の数学②の解答用紙を用い、数字や記号の穴埋めや多肢選択によるマーク式の解答形式である。なお、第 4 問は、M1・M2 冊子の共通問題として出題し、解答欄も数字の穴埋め式となるように揃えられている。

M1・M2 冊子の第 1 問は、2 つの中間〔1〕,〔2〕から構成される。〔1〕は、係数に文字式が入っている方程式の解法(証明)に関する問題である。ただし、解答形式の制約により、M1 冊子(記述式)と M2 冊子(マーク式)で解答させる内容は若干異なる。M1 冊子では、問題文として与えられている証明の中に 2 箇所(2 つの行)の誤りが含まれており、これらの誤りの箇所を指摘するとともに、その修正案を短文で解答させる問題となっている。M2 冊子では、上述した証明の中で行われている場合分けに着目し、それをどのような値によって行うかを多肢選択式で解答させる問題となっている。次の〔2〕は、平成 24 年度数学 I・数学 A (本試験) 第 2 問の(1)までを改題した問題を用い、2 次関数についての知識を問う問題である。また、同一の内容の問題を異なる解答形式(記述式とマーク式)で解答させるため、M2 冊子ではもともとの問題と同じく、数式の一部を穴埋めして数値を解答させる形式となっているのに対し、M1 冊子では数式全体を書かせる解答形式となっている。

M1・M2 冊子の第 2 問と第 3 問は、第 1 問〔2〕と同様に、同一の内容の問題を異なる解答形式(記述式とマーク式)で解答させる問題である。ただし、解答形式の制約により、解答させる内容は両冊子で若干異なる。M1 冊子では、数式や数値を解答させるようになっているが、M2 冊子では、数式の一部を数値と記号を用いて数学②方式により解答させるようになっている。第 2 問は、ある高度を飛行している人工衛星とそれから見える地球上の最も遠い地点との距離を求める問題を題材として扱い、円の接線の方程式や極限の知識を応用して解く問題である。第 3 問は、問題文に示された点列生成のアルゴリズムを理解して解く問題となっており、いくつかの点列を具体的に求めたり、直線と 2 次関数に囲まれる部分の面積を定積分によって求めたりすることな

どが問われている。

M1・M2 冊子の第4問は、M1・M2 冊子の共通問題とし、同一の内容かつ同一の解答欄の問題を異なる解答形式（記述式とマークシート式）で解答させる問題である。カーシェアリングの問題を題材して、確率計算や箱ヒゲ図などの理解を問う問題である。

第1回調査では、部会で作成された問題についての解答データを収集するため、受験者集団を同質とみなせる2群に分割し、M1 冊子と M2 冊子を解答させた。解答終了後にはアンケート調査を実施し、解答方法の説明のわかりやすさ、解答時間が十分だったか、解答用紙のスペースが十分だったか、各大問の難易度に関する評定（以上はいずれも3件法）、および、難しかった点や題意がわかりにくかった点に関する自由記述による回答データを収集した。

第2回モニター調査（以下、第2回調査）では、第1回調査と同様に2種類の問題冊子（数学領域 M3, 数学領域 M4）を準備し、これらに基づく調査が行われた。第2回調査でも、第1回調査と同様に、同一の内容または同様な内容の問題に対して、記述式とマーク式という2種類の解答形式で解答させた場合に受験生の解答状況を比較・検討することを調査目的のひとつとした。また、新たな問題もいくつか作成し、新テストではどのような出題が可能かを検討するための情報を得ることも引き続き行った。M3 冊子と M4 冊子は、表3に示す4つの大問で構成されている。

表 3: 数学の問題冊子（M3 冊子, M4 冊子）の構成

大問 〔中間〕	問題の内容	解答形式	
		M3	M4
第1問	〔1〕 (M3) 証明の間違い探しとその修正 (M4) 証明の場合分け	数式や短文の 記述	数学②方式 (マーク式)
	〔2〕 論理（対偶, 反例, 必要条件, 十分条件, 必要十分条件）	数式や短文の 記述, 選択式	数学②方式 (マーク式)
第2問	「現実の世界」の問題（人工衛星から見た地平線までの距離）への数学的知識（円の接線, 極限）の応用・活用	数式や数値の 記述	数学②方式 (マーク式)
第3問	異なる2つの「現実の世界」の問題（製品の生産個数, リフレッシュ効果の時間数）に共通する数学的定式化（数理計画法）の理解と適用	グラフ, 数式, 数値, 短文の 記述	数学②方式 (マーク式)
第4問	「現実の世界」（カーシェアリング）での問題解決を題材とした確率の理解とデータの解釈	数学②方式 (記述式)	数学②方式 (マーク式)

解答形式は、M1・M2 冊子と同様に、M3・M4 冊子で異なる。M3 冊子は、空欄に数式、数値、短文等を記述させる解答形式に加え、グラフを描かせる解答形式も追加されている。M4 冊子は、センター試験の数学②の解答用紙を用い、数字や記号の穴埋めや多肢選択によるマーク式の解答形式である。

表4に、第1回・第2回調査で用いた問題の対応と変更点をまとめる。M3・M4 冊子に含まれる4つの大問のうち、新規に作成された問題は、M3・M4 冊子のそれぞれにおいて、第1問〔2〕と第3問である。第1問〔1〕は、M3 冊子では、第1回調査で行われた数学領域 M1 に対するア

アンケートの集計結果を考慮し、問題文の体裁の修正がなされたが、M4 冊子では、第 1 回調査と同一の問題が出題されている。第 2 問も、第 1 回調査と同一の問題が使用されている。第 4 問は、4 冊子の共通問題であるが、第 2 回調査では M1・M2 冊子に含まれていたタイプミスの修正が行われて出題されている。

表 4: 第 1 回モニター調査と第 2 回モニター調査の問題の対応

問題番号	M1 冊子から M3 冊子への変更点	M2 冊子から M4 冊子への変更点
第 1 問	[1] 問題文の体裁の修正	変更なし
	[2] 新規作成	新規作成
第 2 問	変更なし	変更なし
第 3 問	新規作成	新規作成
第 4 問	タイプミスの修正	

以下では、M1・M2 冊子と比較して変更された問題の内容や解答形式をまとめる。M3・M4 冊子の第 1 問は、2 つの中間 [1], [2] から構成される。[1], [2] は、それぞれにおいて同様な内容の問題を異なる解答形式で解答させることを意図して作成された問題である。M3 冊子第 1 問 [1] は、M1 冊子 [1] と比較して、問題文の体裁が修正された。これは、第 1 回調査のアンケートで寄せられた意見を反映してのことである。M4 冊子の [1] は特に問題ないと判断され、変更は行われなかった。次に、[2] については、M3・M4 冊子ともに新規の問題を入れることとなった。M4 冊子では、平成 25 年度数学 I・数学 A (本試験) の第 2 問 [2] の論理の問題をそのまま出題することとなった。M3 冊子では、マーク式で解答させるこの問題を記述式に直し、さらに若干の問題を追加して出題することとなった。[2] の前半は、命題の対偶、あるいは命題に対する反例を解答させる問題であり、センター試験と同じ内容であるが、解答形式が異なるものである。一方、[2] の後半は、センター試験の問題をもとに、新傾向の問題を作成した。2 つの条件 p , q の組が与えられている場合に、 p は q であるためのどのような条件であるかについてを、「必要十分条件である」、「必要条件であるが、十分条件ではない」、「十分条件であるが、必要条件ではない」、「必要条件でも十分条件でもない」のいずれかに分類させるという問題である。

M3・M4 冊子第 3 問は、第 1 問と同様に、同様な内容の問題を異なる解答形式で解答させることを意図して作成された問題である。「製品の生産個数」(平成 26 年度工業数理基礎 (本試験) の第 2 問)、「リフレッシュ効果の時間数」という、異なる 2 つの「現実の世界」の問題を題材として、これらの問題に共通する数学的定式化 (数理計画法) の理解と適用を問う内容の問題となっている。M4 冊子では、すべての条件を満たす領域を多肢選択式で選ばせたり、数式の一部を数値と記号を用いて数学②方式により解答させたりするようになっているが、M3 冊子では、グラフを描かせたり、単語や短文を書かせたり、数式や数値を書かせたりするような解答形式になっている。

第 2 回調査では、第 1 回調査と同様に、受験者集団を同質とみなせる 2 群に分割し、M3 冊子と M4 冊子を解答させた。解答終了後にはアンケート調査を実施し、解答方法の説明のわかりやすさ、解答時間が十分だったか、解答用紙のスペースが十分だったか、各大問の難易度に関する評定 (以上はいずれも 3 件法)、さらには、難しかった点、題意がわかりにくかった点、改善してほしい点などに関する自由記述の回答データを収集した。

5.2.2 第1回モニター調査に関する分析

第1回調査の受験者は、平成27年度センター試験「物理」の受験経験を基準として、文系（「物理」（新課程）の受験経験なし）、理系（「物理」（新課程）の受験経験あり）に分類した。

5.2.2.1 M1・M2 冊子全体

各冊子の受験者を合計点の順に得点上位群から下位群まで5群に分割して、各群の得点率をプロットした得点率分析図を作成して、識別力について検討した。冊子全体の合計点の得点率分析図に着目すると、M1・M2冊子ともに、得点下位群から得点上位群までを識別できていることがわかる。しかし、M1とM2の受験者全員の平均正答率（得点率）を比較すると、両者は0.4前後であるが（0.36と0.44）、M1のほうが低い。これは文系受験者が低得点寄りに分布していることに起因している。文系受験者のM1・M2の平均正答率は0.23と0.27、理系受験者のM1・M2の平均正答率は0.44と0.54である。また、M1・M2の合計点の範囲（＝最大値－最小値）に着目すると、受験生全員、文系受験者、理系受験者のそれぞれでM1の合計点の範囲のほうが狭い。合計点の四分位範囲（＝第3四分位数－第1四分位数）に着目すると、受験者集団を文系・理系と分割した場合には、上記の結果とは逆にそれぞれでM1のほうが値が大きくなるが、受験者全員に対してはM2のほうが値が大きい。また、標準偏差に関しては、範囲と同様な傾向が見られ、受験生全員、文系・理系受験者のそれぞれにおいて、M2よりもM1のほうが値が小さいが、いずれも大差はないといってよいであろう。

5.2.2.2 M1・M2 冊子第1問

第1問は、M1・M2全員の得点率は0.53と0.57で大差はなく、得点下位群から得点上位群までの識別は同じようにできている。しかし、文系受験者と理系受験者の得点率には大きな開きがあった。M1では文理の得点率の差は0.26（文系：0.37、理系：0.63）、M2の文理の得点率の差は0.33（文系0.37、理系0.70）である。M1冊子〔1〕の証明の誤りを2箇所指摘する問題（解答記号ア、イ）は、全体的には、片方を誤答するケースがいくつも見られた。M2冊子〔1〕の証明の場合分けの値を解答する問題について、全員の正答率は、解を一意に求められる場合が一番低く（0.14）、不能と不定の場合の正答率はこれよりも高く、それぞれ0.62と0.48であった。M1・M2冊子〔2〕の2次関数に関する知識を問う問題は、3つの採点単位（設問）があるが、いずれもM1（記述式）のほうが低かった（0.66と0.79、0.51と0.70、0.28と0.59）。得点率の低い3つ目の採点単位は、無解答も多く、M1・M2のそれぞれで全体の10%以上となっていた。

5.2.2.3 M1・M2 冊子第2問

第2問は、現在行われているセンター試験の数学の問題と比較して、誘導を極力減らし、より深い思考力を問うことを意図して作成された問題であり、4つの大問の中で得点率は一番低かった（M1全員：0.14、M2全員：0.34）。M1・M2の得点上位群の正答率は、それぞれ0.32と0.62であった。M1の5つの採点単位のうち、最初の4つは数式を書かせる問題であり、最後は計算により数値を解答させる問題である。対応するM2では、最初の4つの採点単位は数式の係数を穴埋めする形式で、最後はM1と同様に計算により4桁の解答欄に数値を解答させる問題である。M1・M2の5つの採点単位のそれぞれで正答率を比較すると、出題順に、0.58と0.78、0.14と

0.42, 0.18 と 0.67, 0.12 と 0.23, 0.01 と 0.12 であり, また全体的に無解答の多い (10%以上) 問題であった。

5.2.2.4 M1・M2 冊子第 3 問

第 3 問は, 現在行われているセンター試験の数学の問題と比較して, 問題文の分量の多い問題で, 点列生成のアルゴリズムを理解して解答する問題である。この大問は, 4つの大問中 2 番目に正答率が低く, M1・M2 のそれぞれの全員の正答率は, 0.25 と 0.30 であった。採点単位は, M1・M2 冊子とも全部で 8 つあるが, この大問に関しては, いずれの採点単位においても M2 冊子の正答率のほうが高かったが, M1・M2 冊子間で正答率の差は最大で 0.11 であった。この大問は, 3 つ目の採点単位以降, 無解答が 10%以上となる問題であった。

5.2.2.5 M1・M2 冊子第 4 問

第 4 問は, 同一の内容の問題が出題され, 解答形式が記述式かマーク式かの違いだけのある問題であった。M1・M2 冊子ともこの大問の得点率は 0.52 であり, 大問全体としては差は見られなかった。ただし, 各設問の正答率には差があり, 9 つの採点単位のうち, 最も得点率が開いたのは最後の設問の 0.13 で, この設問に関しては M1 のほうが M2 の得点率を上回った。

5.2.3 第 2 回モニター調査に関する分析

第 2 回調査の受験者は, 平成 27 年度センター試験の「地理歴史・公民」や「理科」で選択解答した科目数を基準として, 文系 (「地歴・公民」2 科目を選択解答した者), 理系 (理科②から 2 科目, または, 理科①と理科②の両方を選択解答した者) に分類した。

5.2.3.1 M3・M4 冊子全体

冊子全体の合計点の得点率分析図に着目すると, M3・M4 冊子ともに, 得点下位群から得点上位群までを識別できている。しかし, M3 と M4 の全員の平均正答率 (得点率) を比較すると, M1・M2 の比較と同様に, 記述式の M3 のほうが低く (M3:0.62, M4:0.68), 文系受験者が低得点寄りに分布している。M3・M4 の文系受験者の平均正答率は 0.52 と 0.58, 理系受験者の平均正答率は 0.68 と 0.75 である。また, M3・M4 の合計点の範囲 (= 最大値 - 最小値) に着目すると, 受験生全員, 文系受験者, 理系受験者のそれぞれで M3 の合計点の範囲のほうが広い。合計点の四分位範囲 (= 第 3 四分位数 - 第 1 四分位数) に着目すると, 受験者集団を文系・理系と分割した場合には, 理系受験者は M3 のほうが値が大きい, 文系受験者あるいは受験者全員に対しては M4 のほうが値が大きい。また, 標準偏差に関しても四分位範囲と同様な傾向が見られるものの, M3・M4 間で大差はないといってよいであろう。

5.2.3.2 M3・M4 冊子第 1 問

第 1 問は, M3・M4 全員の得点率は 0.74 と 0.67 で, 得点下位群から得点上位群までの識別は同じようにできている。また, M3 と M4 のそれぞれで, 文系・理系受験者間の得点率の差も同程度であった。M3 冊子 [1] の証明の誤りを 2 箇所指摘する問題は, 片方を誤答する者もいたが, 全体的にはよくできていた。M4 冊子 [1] の証明の場合分けの値を解答する問題について, 全員の正答率は, 解を一意に求められる場合が一番低く (0.23), 不能と不定の場合の正答率はこれよ

りも高く、それぞれ0.82と0.65であった。この傾向はM2冊子と同様であるが、得点率はM2よりもM4のほうが高い。この中間はM2・M4冊子で同一の問題であるので、得点率の違いは、受験者層の違いを反映しているといつてよいであろう。M3・M4冊子〔2〕の論理に関する問題は、最初の3つの採点単位（設問）は、同一の内容の問題を異なる解答形式で解答させる問題である。しかし、いずれもM3（記述式）のほうが低かった（0.84と0.93、0.85と0.90、0.56と0.78）。

5.2.3.3 M3・M4冊子第2問

M3・M4冊子第2問は、M1・M2冊子第2問と同一の問題である。誘導を極力減らし、思考力を問うことを意図して作成された問題であり、4つの大問の中で得点率は一番低かった（M3全員：0.35、M4全員：0.52）が、M1・M2第2問の得点率と比較すると、いずれも0.2ほど高く、これもM1・M2の受験者層とM3・M4の受験者層の違いを反映しているものと考えられよう。M3の5つの採点単位のうち、最初の4つは数式を書かせる問題であり、最後は計算により数値を解答させる問題である。対応するM4では、最初の4つの採点単位は数式の係数を穴埋めする形式で、最後はM3と同様に計算により4桁の解答欄に数値を解答させる問題である。M3・M4の5つの採点単位のそれぞれで正答率を比較すると、出題順に、0.79と0.91、0.37と0.71、0.49と0.82、0.38と0.52、0.16と0.26であり、M3では2つ目以降の採点単位で、M4では3つ目以降の採点単位で、それぞれ無解答が10%以上となる問題であった。

5.2.3.4 M3・M4冊子第3問

第3問は、M1・M2冊子第3問と同様に、現在行われているセンター試験の数学の問題と比較して、問題文の分量の多い問題である。またこの大問は、異なる2つの「現実の世界」の問題を題材として、それらに共通する数学的定式化（数理計画法）を理解して解答する、新傾向の問題である。この大問は、M3冊子では4つの大問中2番目に正答率が低かったが、M4冊子では最も得点率の高い大問であった。採点単位は、M3冊子は7つ、M4冊子は5つあり、両者の採点単位数は異なっているものの、解答形式の違いにより正答率が比較可能と考えられるものは、M3の採点単位のうち、1、4～7番目の5つである。それぞれについてのM3・M4の正答率は、0.74と0.89、0.60と0.87、0.78と0.93、0.48と0.67、0.45と0.65であり、いずれも0.2～0.3程度M3の正答率のほうが低い。この大問は、M3では3つ目以降の採点単位で、M4では4つ目以降の採点単位で、それぞれ無解答が10%以上であった。

5.2.3.5 M3・M4冊子第4問

M3・M4冊子第4問は、M1～M4冊子の共通問題として出題され、M3・M4の違いは、解答形式が記述式かマーク式かだけである。M3・M4冊子の各大問の得点率は、0.76と0.75であり、大問全体としては大差はなく、M1・M2冊子の各大問得点率（いずれも0.52）と比較すると、0.2ほど高かった。この傾向は、第2問と同様であり、受験者層の違いによるものと考えられよう。ただし、各設問の正答率には差があり、9つの採点単位のうち、最も得点率が開いたのは7番目の設問の0.12である。

5.2.4 まとめ

数学部会では、記述式とマーク式で解答させる問題冊子を2種類ずつ作成し、2回のモニター調査で856名の大学1年生にこれらを解答してもらうことにより、新テストでどのような問題が出題可能であるかを検討するための情報を収集した。記述式の採点では、調査実施前に作成した採点基準を、調査実施後に見直して再検討する必要があることを確認し、例えば、問題作成時に想定しなかった別解を追加したり、正答率が極端に低い問題については、答案を見てから正答の許容範囲を広げたり、といった検討を行った。

今回の調査で使用した問題は、どちらの解答形式でも、冊子全体および各大問の識別力があつた。すなわち、得点上位群になるほど正答率が高くなり、得点下位群になるほど正答率が低くなる傾向が見られた。ただし、難しい大問では0点付近あるいは低得点寄りに分布が集中するため、下位群の識別が鈍くなる。

2種類の解答形式で得られた冊子全体の合計点や大問得点の得点率、あるいは設問の正答率を比較すると、記述式のほうがマーク式よりも正答率（得点率）が低くなる傾向が見られた。特に記述式においては、受験者集団を文系受験者と理系受験者に分けた場合、文系受験者は理系受験者と比較して、0点付近または低得点寄りに分布が集中する傾向が見られた。しかし、合計点の標準偏差、四分位範囲に着目すると、記述式とマーク式の間でそれぞれの値に大差は見られなかった。

5.3 世界史

5.3.1 問題冊子の構成（第1回、第2回）

第1回モニター調査で用いたWH冊子は、歴史的思考力や表現力を測定することを目的として、表1に示す4つの大問で構成されている。選択式問題や並べ替え問題は、歴史資料をよみとき、複数の歴史事象を関連付けながら、多面的・多角的に考察する必要がある問題である。また記述式問題は、歴史資料をよみといた上で、歴史事象に関する基本的な理解に基づいて適切に記述する問題である。

表 1: 世界史の問題冊子 (WH 冊子) の構成

大問	題材	設問	解答形式
第1問	3地域における体制の変化を示す略地図	問1～問6	選択式
		問7	記述式（1行）
第2問	アフリカにおける鉄道建設の発展を描いた地図	問1	並べ替え
		問2	選択式
第3問	第1次世界大戦中に制作されたポスターとそれに関する会話	問1～問3	選択式
		問4	40字の記述式
第4問	ドイツにおける庇護権申請数に関するグラフとそれに関する記事	問1, 問2	選択式
		問3	記述式（1行）

第2回モニター調査で用いたWH2冊子も、歴史的思考力や表現力を測定することを目的として、表2に示す6つの大問で構成されている。

表 2: 世界史の問題冊子 (WH2 冊子) の構成

大問	題材	設問	解答形式
第1問	4地域の経済規模に関する折れ線グラフとそれに関する会話	問1～問6	選択式
第2問	3地域における体制の変化を示す略地図	問1～問6	選択式
		問7	4字
第3問	アフリカにおける鉄道建設の発展を描いた地図	問1	並べ替え
		問2	選択式
第4問	第1次世界大戦中に制作されたポスターとそれに関する会話	問1～問3	選択式
		問4	約40字の記述式
第5問	日本の東南アジアへの経済進出の政治・外交上の背景		約100字の記述式
第6問	ドイツにおける庇護権申請数に関するグラフとそれに関する記事	問1, 問2	選択式
		問3	約15字の記述式

WH冊子と同様に、選択式問題や並べ替え問題は、歴史資料をよみとき、複数の歴史事象を関連付けながら、多面的・多角的に考察する必要がある問題である。また記述式問題は、歴史資料

をよみといた上で、歴史事象に関する基本的な理解に基づいて適切に記述する問題である。6つの大問のうち、第2問、第3問、第4問、第6問は、第1回モニター調査で試行した問題をもとに、問いたい内容がより明確になるように設問の一部に変更が加えられている。第1問と第5問は新規の問題である。第1問は、歴史資料とともに教師と生徒の会話に関して論理的に思考する必要のある問題であり、第5問は、複数の時代の歴史事象について多面的に比較・考察し、記述する問題である。

WH2冊子の多肢選択式の設問のうち、第1問の問1と問2では、選択肢から正しいものをすべて選ぶ問題である。第1問の問3では、組合せを1つ解答するよう指示されている問題であるが、正しい組合せが複数ある問題である。

WH冊子、WH2冊子共に、解答時間は70分を想定して作成されている。各冊子の各設問には配点が付けられており、WH冊子の満点は100点、WH2冊子の満点は150点である。

第1回モニター調査では、解答にかかる時間を調査するため、試験時間中10分ごとに、その時点で解答している問題に該当する解答用紙の所定欄に○を付けるよう、監督者が指示した。第1回モニター調査において、約4分の3の受験者が30分経過の時点で解答を終えており、設問だけを見て解答を諦める傾向が示唆されたため、第2回モニター調査では、大問の数を増やすと共に、リード文を工夫して、受験者に資料を読ませるよう誘導を図った。

いずれの調査でも、解答終了後のアンケートで、各設問に解答する際に最も使用したと思う能力（「知識」「資料読解力」「思考力」「表現力」）や（「表現力」は第2回モニター調査のみ）、センター試験と比べた印象などを回答させた。

5.3.2 第1回モニター調査に関する分析

WH冊子の各設問の配点は5点あるいは10点であり、100点満点である。表3に、合計点及び大問ごとの基本統計量を示す。「得点率」とは、各大問の満点に対する大問得点の平均点の割合である。合計点の平均点は約42点であり、各大問の得点率も4割～5割である。

受験者217名を世界史の合計点の順に、得点の高い群（H群）、やや高い群（MH群）、やや低い群（ML群）、低い群（L群）の四つに分割して、各設問の得点率（正答率）を算出した結果を、表4に示す。記述式や並べ替えのように、正答か誤答に分けられない設問については、その設問の配点（満点）に対するその設問の平均点の割合を、得点率としている。

得点率が4割～6割の設問が多く、ほとんどの設問で、L群からH群になるにつれて得点率が上がる傾向がみられる。記述式の設問には、第1問の問7（全体の得点率0.10）、第4問の問3（全体の得点率0.24）のように、正答率の極端に低い問題がみられる。記述式の設問では、15～20字程度であっても、正しい内容の他に誤りを含む内容も書いてしまう答案が目立つ。資料を読み解いて五つの出来事の順番を時代順に並べ替える設問（第2問の問1）では、得点率は35%であるが、五つすべての順番を正しく並べることができた者の割合は7.8%に過ぎない。アンケート調査の結果をみると、知識を問う設問と受験生に捉えられてしまっており、資料を読みとこうとしない傾向が示唆されている。

5.3.3 第2回モニター調査に関する分析

WH2冊子は150点満点である。表5に、合計点及び大問ごとの基本統計量を示す。「得点率」とは、各大問の満点に対する大問得点の平均点の割合である。合計点の平均点は約68点、得点率

表 3: 基本統計量 (WH 冊子)

	合計点	第 1 問	第 2 問	第 3 問	第 4 問
受験者数	217	217	217	217	217
満点	100	40	15	25	20
平均	41.84	15.97	7.76	10.58	8.55
得点率	0.42	0.4	0.52	0.42	0.43
標準偏差	14.15	7.45	4.65	6.3	5.33
最小値	15	0	0	0	0
最大値	85	30	20	25	20

表 4: 各設問の得点率 (全体及び 4 群) (WH 冊子)

	全体	L 群	ML 群	MH 群	H 群
第 1 問 問 1	0.47	0.19	0.30	0.55	0.65
第 1 問 問 2	0.47	0.28	0.30	0.43	0.74
第 1 問 問 3	0.37	0.19	0.33	0.38	0.47
第 1 問 問 4	0.56	0.44	0.42	0.59	0.69
第 1 問 問 5	0.65	0.44	0.49	0.68	0.86
第 1 問 問 6	0.47	0.13	0.35	0.41	0.76
第 1 問 問 7	0.10	0.02	0.05	0.07	0.20
第 2 問 問 1	0.35	0.20	0.30	0.38	0.44
第 2 問 問 2	0.65	0.41	0.61	0.68	0.75
第 3 問 問 1	0.48	0.19	0.53	0.55	0.51
第 3 問 問 2	0.37	0.25	0.19	0.34	0.60
第 3 問 問 3	0.48	0.28	0.37	0.46	0.68
第 3 問 問 4	0.39	0.05	0.30	0.36	0.64
第 4 問 問 1	0.65	0.28	0.61	0.73	0.76
第 4 問 問 2	0.59	0.41	0.49	0.64	0.71
第 4 問 問 3	0.24	0.05	0.13	0.21	0.43

では45%である。各大問の得点率は2割～6割である。

受験者110名を世界史の合計点の順に、得点の高い群（H群）、中程度の群（M群）、低い群（L群）の三つに分割して、各設問の得点率（正答率）を算出した結果を、表6に示す。

記述式や並べ替えのように、正答か誤答に分けられない設問については、その設問の配点（満点）に対するその設問の平均点の割合を、得点率としている。設問の得点率は約1割～8割と幅広い。複数の正答がある設問（第1問の間2、第1問の間3）や、記述式の設問（第5問、第6問の間3）での得点率が低い傾向がみられる。資料を読み解いて五つの出来事の順番を時代順に並べ替える設問（第3問の間1）について、WH2冊子では部分点は与えず、五つすべての順番が正しいものだけを正答としたが、正答率は22%であり、WH冊子で五つすべての順番を正しく並べることができた者の割合7.8%に比べて上昇している。これは、資料に注目させるようにリード文を変更した効果と思われる。

第1問の間1は、6つの選択肢から正しいものをすべて選ぶ設問である。正答選択肢は②、④、⑥の3つである。表7に、第1問の間1で各選択肢を選択した人数とその割合を全体及び群ごとに算出して示す。表8に、選択した選択肢の組合せと、その選択率を示す。正答選択肢②、④、⑥それぞれの選択率は94%、57%、69%と高いが、3つすべてを過不足なく選択できた受験者は全体の23%と少ない。同様の傾向は、第1問の間2にもみられる。第1問の間2は、8つの選択肢から正しいものをすべて選ぶ問題である。2つの正答選択肢それぞれは、56%、37%の選択率であったが、正答選択肢2つを過不足なく選択できた受験者は2%に過ぎない。

5.3.4 まとめ

新傾向として、問題中で提示した歴史資料をよみとき、複数の歴史事象を関連付けながら、多面的・多角的に考察する必要がある問題を出題した。記述式の設問は、歴史資料をよみといた上で、歴史事象に関する基本的な理解に基づいて記述するもので、15～40字程度の記述が中心であるが、第2回モニター調査では100字程度の記述も出題した。記述式の設問の得点率は、いずれの冊子でも0.10から0.39と低かった。

WH2冊子（第2回モニター調査）では、正答選択肢をすべて選ぶ設問を何題か出題したが、いずれも、正答選択肢を過不足なくすべて選べた受験生の割合は非常に低かった。今回の調査では、正答選択肢については選んだ数に応じて加点し、誤答選択肢については選んだ数だけ減点する（マイナス点にはしない）方式で採点を行ったが、このようなタイプの設問では、どこまで分かっているかが適切に得点に反映されるよう、部分点の与え方について検討する必要がある。

資料を読み解いて五つの出来事の順番を時代順に並べ替える設問は、WH冊子（第1回モニター調査）ではすべてを正しく並べることができた者の割合は7.8%と低かった。資料を読み解き、基本的な知識と合わせて考えれば正しく並べることができる設問であったが、アンケート調査の結果をみると、知識のみを問う設問と受験生に捉えられて、受験生が資料を読みとこうとしない傾向が示唆されたため、WH2冊子（第2回モニター調査）で、資料に注目させるようにリード文を変更した。WH2冊子（第2回モニター調査）ではすべてを正しく並べることができた者の割合は22%にまで増えた。アンケート調査でも、「資料読解力」を問われていると捉えた回答が多くなった。

表 5: 基本統計量 (WH2 冊子)

	合計点	第 1 問	第 2 問	第 3 問	第 4 問	第 5 問	第 6 問
受験者数	110	110	110	110	110	110	110
満点	150	30	40	15	25	20	20
平均	67.92	13.44	22.55	5.09	12.00	5.12	12.00
得点率	0.45	0.45	0.56	0.34	0.48	0.26	0.60
標準偏差	20.92	4.00	10.55	5.09	6.44	5.39	6.44
最小値	22	5	0	0	0	0	0
最大値	116	24	40	15	25	20	25

表 6: 各設問の得点率 (全体及び 3 群) (WH2 冊子)

	全体	L 群	M 群	H 群
第 1 問 問 1	0.58	0.41	0.55	0.76
第 1 問 問 2	0.09	0.04	0.07	0.16
第 1 問 問 3	0.10	0.03	0.10	0.15
第 1 問 問 4	0.16	0.14	0.17	0.18
第 1 問 問 5	0.97	0.97	0.94	1.00
第 1 問 問 6	0.69	0.47	0.78	0.82
第 2 問 問 1	0.67	0.50	0.61	0.89
第 2 問 問 2	0.65	0.44	0.67	0.84
第 2 問 問 3	0.48	0.19	0.42	0.82
第 2 問 問 4	0.72	0.53	0.67	0.95
第 2 問 問 5	0.76	0.44	0.86	0.97
第 2 問 問 6	0.58	0.22	0.67	0.84
第 2 問 問 7	0.32	0.06	0.22	0.66
第 3 問 問 1	0.22	0.00	0.17	0.47
第 3 問 問 2	0.58	0.42	0.67	0.66
第 4 問 問 1	0.49	0.50	0.50	0.47
第 4 問 問 2	0.58	0.39	0.56	0.79
第 4 問 問 3	0.64	0.33	0.69	0.87
第 4 問 問 4	0.35	0.18	0.39	0.46
第 5 問	0.26	0.11	0.22	0.43
第 6 問 問 1	0.83	0.75	0.81	0.92
第 6 問 問 2	0.54	0.47	0.64	0.50
第 6 問 問 3	0.29	0.25	0.33	0.29

表 7: 第 1 問 問 1 各選択肢を選択した人数 (全体及び 3 群)

	全体		L 群		M 群		H 群	
	人数	%	人数	%	人数	%	人数	%
1	5	4.55	5	13.89	0	0.00	0	0.00
2*	103	93.64	32	88.89	34	94.44	37	97.37
3	9	8.18	4	11.11	3	8.33	2	5.26
4*	63	57.27	22	61.11	18	50.00	23	60.53
5	42	38.18	18	50.00	17	47.22	7	18.42
6*	76	69.09	14	38.89	26	72.22	36	94.74

表 8: 第 1 問 問 1 (全体及び 3 群)

選択肢	全体	L 群	M 群	H 群
246*	0.23	0.08	0.17	0.42
26	0.19	0.08	0.19	0.29
256	0.14	0.11	0.22	0.08
24	0.10	0.17	0.11	0.03
245	0.10	0.17	0.11	0.03
2456	0.06	0.06	0.06	0.08
236	0.03	0.03	0.03	0.03
25	0.03	0.06	0.03	0.00
2346	0.02	0.00	0.03	0.03
4	0.02	0.06	0.00	0.00

5.4 物理

5.4.1 問題冊子の構成（第 1 回，第 2 回）

物理の問題冊子は，知識・技能を応用する問題を用い，従来の共通試験問題に比べ，判断力・思考力・表現力に関する判定機能を強化することを目標として作成された。表 1 および表 2 にそれぞれ第 1 回モニター調査および第 2 回モニター調査の問題冊子の構成を示す。第 2 回モニター調査の問題冊子は，第 1 回モニター調査で用いた問題を元に，問題文および問題設定の細部の修正を行うとともに，一部の問題の差し替えおよび追加等を行うことにより，新たな数式記述問題，文章記述問題，作図問題を組み込んだ形で構成されている。

第 1 回モニター調査では，一部の問題で同一の内容で選択式と数式記述式の二通りの解答形式を用意し，P1 冊子と P2 冊子に分けて配置した。第 2 回モニター調査では，一部の問題で同一の内容で選択式と記述式（数式記述または文章記述）の二通りの解答形式を用意し，P3 冊子と P4 冊子に分けて配置した。

各モニター調査で用いる 2 つの冊子の間で問題量になるべく偏りが生じないようにし，解答時間が概ね均等になるよう配慮してある。（解答時間 80 分）

表 1: 物理の第 1 回モニター調査の問題冊子の構成（P1，P2 冊子）

大問	題材	設問		解答形式	
		P1	P2	P1	P2
第 1 問	地球に届く太陽光のエネルギーに関する実験および探究	問 1		選択式	文章記述（25 字以内）
		問 2		選択式（共通）	
		問 3		選択式（共通）	
		問 4		選択式（共通）	
第 2 問	アトウッドの装置による物体の運動の把握と推測	問 1		数式記述	選択式
		問 2		選択式（共通）	
		問 3～4		数式記述	選択式
第 3 問	ジェットコースターの運動の把握とデータの読取り	問 1		選択式（共通）	
		問 2		選択式（共通）	
第 4 問	バンジージャンプの力学的解析	問 1		文章記述（150 字以内）（共通）	
		問 2		文章記述（200 字以内）（共通）	
		問 3～5		選択式（共通）	

表 2: 物理の第 2 回モニター調査の問題冊子の構成 (P3, P4 冊子)

大問	題材	設問		解答形式	
		P3	P4	P3	P4
第 1 問	地球に届く太陽光のエネルギーに関する実験および探究	問 1		選択式	文章記述 (25 字以内)
		問 2		選択式 (共通)	
		問 3		数式記述 (共通)	
		問 4~5		選択式 (共通)	
第 2 問	アトウッドの装置による物体の運動の把握と推測	問 1		数式記述	選択式
		問 2		選択式 (共通)	
		問 3~4		数式記述	選択式
第 3 問	ジェットコースターの運動の把握とデータの読取り	問 1		選択式 (共通)	
		問 2	問 2	選択式	作図
		問 3		文章記述 (15 字程度)	
第 4 問	バンジージャンプの力学的解析	問 1		文章記述 (150 字以内) (共通)	
		問 2		数式記述 (共通)	
		問 3~4		選択式 (共通)	

5.4.2 モニター調査の分析結果

本稿では、第 1 回モニター調査 (P1 冊子, P2 冊子) についての分析結果を述べる。第 2 回モニター調査 (P3 冊子と P4 冊子) については、第 1 回モニター調査で得た知見をもとに、採点基準などを精査している段階である。

なお、5.4.2.2 以降で述べる文章記述または数式記述の問題の採点は、一般の受験予備校の物理の受験指導を専門とする立場の者が行った。採点作業は問題作成者が決めた採点基準に基づいて作成された採点指示書に従って行われた。また、各答案を 2 名の者が互いに独立に採点することにより、採点結果の一致・不一致を評価できるようにした。

5.4.2.1 選択式部分について

表 3 と表 4 に、それぞれ、P1 冊子と P2 冊子の選択式部分の合計点の基本統計量を示す。

表 3: 合計点の基本統計量 (P1 冊子)

受験者数	満点 (採点項目数)	平均	平均正答率	標準偏差	最大値	最小値
132	22	12.37	0.56	3.28	20	5

表 4: 合計点の基本統計量 (P2 冊子)

受験者数	満点 (採点項目数)	平均	平均正答率	標準偏差	最大値	最小値
136	28	15.10	0.54	4.19	25	0

5.4.2.2 記述式問題の採点に関する分析 (P2 冊子第 1 問-問 1)

データを読み取る方法の科学的根拠について、物理で学んだ熱の知識に基づいて、思考し判断する能力を問う記述式の問題である。指定文字数は 25 文字以内である。

試験後に実際の答案すべてに目を通して、カテゴリー分けを行い、採点指示書を作成した。実際の採点指示書は、本総合報告書とは別に、テーマ別のより詳細な内容を記載した報告書「新テスト試作問題の作成と評価」の方に収録した。

表 5 に、採点指示書に基づいて 2 名が採点を行った結果を示す。25 文字以内の短文の記述式問題であり、しかも試験後に答案を分析したうえで精緻な採点指示書を作成したにもかかわらず、「採点者によって正誤が異なっていた事例」が 4.4%あった。また、「正誤の判断は一致しているが、カテゴリーの判断が異なっていた事例」が 11.8%あった。仮に、受験者を 10 万人と仮定すると、その人数は 4,400 人となる。ただし、実際に新しい試験が実施される際には、第 2 次的、第 3 次的な採点確認の過程が導入される可能性があり、第 1 次的な採点だけで結果が決まるとは限らない。

表 5

度数	%	一致度	判定
114	83.8%	2 名とも、カテゴリーが一致している。	OK
10	7.4%	カテゴリーは異なるが、2 名とも正解と判断している。	
6	4.4%	カテゴリーは異なるが、2 名とも誤答と判断している。	
6	4.4%	2 名で正誤が異なる。	NG

計 136 計 100%

5.4.2.3 新しい記述方式に関する分析 (P1 冊子第 2 問-問 1, 問 3, 問 4)

P1 冊子, P2 冊子の第 2 問は、アトウッドの装置を題材として、教科書レベルの問いから、高難度大学の志望者を想定受験者とした問いまで、広い学力層に対して識別力を持つことを目指した小問で構成されている。同じ内容の問いを、P2 冊子では全問多肢選択式で、P1 冊子では一部の問いを以下のような新しい数式記述方式に変えて作題した。

センター試験では、数式群を選択肢とする出題が多用されてきた。これにより、正しい式を導くための思考力を測ることが可能となるが、完全ではない。選択肢の存在そのものがヒントとなり、ゼロから数式を組み立てる創造的思考力の測定には限界があるほか、選択肢数に限りがあるため、高難度の問いでの測定誤差が大きくなる。ゼロからの創造的思考力を測定することを目的としながら、採点に多大な人・日を要する通常の記述式よりは簡便な解答方式として、数式を計算機処理により採点できる形に記述させる、新しい数式記述法 (以下、新記述法と呼ぶ) を試用することとした。

新記述法の規則は「解答上の注意」として、問題冊子の裏表紙に次のような箇条書きとして提示した。

数式の記述法

- ・左詰めで、1つの枠に1つの文字を記入し、記述の最後には、終了記号（#）を書き込む。
- ・使用できる文字は、数字（0～9）、大文字および小文字の英字（A～Z およびa～z）、四則演算子（+、-、*、/）、べき乗演算子（ \wedge ）、左および右かっこ（（ および ））、ピリオド（.）、コンマ（,）である。
- ・演算の優先順位は、べき乗→乗除算→加減算の順とし、乗算は星印（*）を、除算は斜線（/）を用いる。
- ・乗算の記号（*）は省略できる。
- ・優先順序を明示するため、必要に応じて括弧でくくる。左右の括弧は対になっていなければならない。
- ・関数の独立変数は括弧でくくる。 $\sin x$ は $\sin(x)$ と書く。
- ・平方根は $\text{sqrt}(x)$ と書く。
- ・記述の最後には、終了記号（#）を書き込む。

今回の調査では、そのフィージビリティを調べるため、P1冊子の受験者132名に対して、新記述法による解答に併せて通常の記述法による解答も求め、2名の採点者により、目視による採点を行った。（本問題の採点に用いた採点指示書についても、テーマ別の報告書「新テスト試作問題の作成と評価」の方に収録した。）

採点指示書を通じて、数式を求めさせる7つの設問（解答番号1, 2, 3, 4, 11, 12, 13）のそれぞれについて、解答結果を、通常の記述法では4つ、新記述法では7つのグループのいずれか1つに分類させた（表6）。通常の記述法のグループ3（指定外の文字）は解答番号3・4のみが該当し、問題文で指定された文字以外の使用を意味する。新記述法のグループ3（新記述法で正しく再現）は空白・正解（グループ1・2）以外の通常の記述法の解答が新記述法に従った形で正しく再現されていること、グループ4は末尾に#を補えば、上の意味で、正しく再現されている

表6 解答のグループ分け

グループ	通常の記述法	新記述法
1	空白	空白
2	正解	正解
3	指定外の文字	新記述法で正しく再現
4	その他の不正解	末尾の#が欠落
5		指定外の文字・記号
6		括弧が不足
7		その他の不正解

こと、グループ 5（指定外の文字・記号）は裏表紙に書かれた新記述法の規則違反、グループ 6 は計算順序を指定する括弧を補えば、正しく再現されていることを意味する。

採点（分類）作業は 1 枚の答案を 2 名の採点者が担当した。両者の判定結果が一致するものはそのままそれを結果とし、2 名の結果が不一致であったものは、スキャンされた原本にあたって判定し直すこととした。

2 名の採点結果の不一致は、924 件のサンプルのうち、通常の記述法で 32 件、新記述法で 75 件が不一致であった。このうち通常の記述法では、論理的にあり得ない分類（解答番号 3・4 以外の解答についてグループ 3）が 20 件あり、本質的な不一致は 12 件であった。通常の記述法では、本質的な不一致の 3 分の 2 の 8 件が正解・不正解の判定に関わるものであった。新記述法では、不一致の大部分が不正解の理由の判定に関するものであり、残る 17 件が正解・不正解の判定に関わるものであった。

通常の記述法での正答率は、教科書の練習問題レベルの解答番号 1・2 で 6～7 割、その簡単な応用問題である解答番号 3・4 で 5 割前後、高難度の解答番号 11・12・13 で 1 割以下であった。

通常の記述法で書かれた正解のうち、新記述法で正しく再現できたものは概ね半数であった。事前の訓練が全くなかった状態で、新しい規則を理解してそれに従って数式を記述するという課題を考えると、妥当な再現率であろう。再現できなかった場合の最大（6 割弱）の要因は、電子化する際に、消し跡などにより記述の終わりの判定が困難になるのを避けるために導入した、末尾に # を書くという規則を守れなかったことであった。この規則は CBT で受験者がキー入力する場合には必要なくなる。今回の結果で、# 欠落を容認することにすれば、正解の再現率は、4 分の 3 まで上がる。

次に多かったの（約 2 割）が分数などで、計算の順序を指定する括弧を正しく用いていなかったものである。これは、分数を斜線を用いて 1 行に記述する例が高校教科書にはほとんどなく、日常訓練される機会がないためでもあろう。

今回は初めての試みであったので、得られた正解を、正しく再現する割合はやや低かったが、事前に周知するなどすれば、各教科科目で許容されうる解答方法を用いることも検討の余地があると考えられる。

5.4.2.4 正答が 2 つある出題に関する分析（P1 冊子・P2 冊子第 3 問）

ジェットコースターのような移動物体の運動の解析においては、変位・速度・加速度に関する事項が扱われる。これらは、物理基礎の学習内容としては、学年での学習の前半に学ばれるもので、これらの学習内容に関しては、受験生は比較的習熟しているのではないかと考えられる。そのため、これまでの大学入試センターの出題問題としても扱われてきた題材である。

これまでの大学入試センターでの出題形式においては、1 つの正解のみを問う出題形式であったため、選択肢を見ながらだいたいの検討をつければ、あとは、消去法などの方法で正答にたどり着くことも可能であり、本質的に理解をしているので正答しているのか、選択肢の選択がうまくいったので本質的な理解に届いていなくても正答に至ったのかの判断がつきにくいという問題点が残されてきた。そこで、今回は、選択肢のなかに複数、本問の場合は 2 つの正答があるよう

な出題を行い、これまでの大学入試センターにおける物理の問題がもっていた問題点の改善ができるかどうかを検証した。

第3問は、P1冊子とP2冊子の両方で同じ問題を出題した。全体としての平均正答率は45%程度であった。受験者が、高校で選択物理を受講してきた大学生のデータであることから判断すると、複数選択にすることによって、難易度があがったといってもよいといえよう。

5.4.2.5 記述問題に関する分析 (P1冊子・P2冊子第4問-問1, 問2)

バンジージャンプの背後にある物理的原理は、一定の大きさの重力と大きさが変化するばねによる弾性力という性質の異なる2つの力が競い合う興味ある力学系となっており、力学系での問題として重要な地位を占める。このような力学系の振る舞いを深く理解するには、各々の状況における力、運動方程式、力学的エネルギーの相互関係を良く思考し、それらを相互に関係づけて理解できる力が重要である。本問題は、受験生がこの様な状況を的確に把握し、その筋道を論理的に述べ、論理に従って数式化を行い、解答を得るプロセスを実行できるか否かを確かめるため、試行的に出題された。

問1と問2は、考える筋道を論理的に記述できるかを測る目的で出題されたもので、問題解決への道筋を、自分の考えに沿って科学的文章として論理的に記述する能力をチェックするための出題である。

問1では、運動方程式と力の関係を記述させた(150字)。採点時の大きな混乱を防ぐため、「下向きの重力」と「上向きのばねの弾性力」という語句を文中に取り入れることを要請し、文章の自由度を制限した。問2では、運動方程式と相補関係にある力学的エネルギー保存の法則を記述させた(200字)。採点時の大きな混乱を防ぐため、「重力による位置エネルギー」「ばねの弾性力による位置エネルギー」「運動エネルギー」という句を文中に取り入れることを要請し、文章の自由度を制限した。

これらの記述式問題の採点では、解答文章の中の要素別にその要素が満たされるか否かを採点者に問い、フローチャート式に採点させた。また、その判断の確信度についても記入してもらった。(本問題の採点に用いた採点指示書についても、テーマ別の報告書「新テスト試作問題の作成と評価」の方に収録した。)その結果、採点基準がそれなりに機能することが判明した。これらの問いでは、文中に挿入すべき語句を指定したこともあり、全体として採点は予期した程困難ではなかったが、それでも問1では約5%、問2では約2%の採点不能答案が出た。問2で採点不能答案が少ないのは、この問いを放棄した受験生が多かったことによると考えられる。採点不能答案は、解答文全体が論理的でなく、部分的に正しいことが書かれているような場合で、例え出題者が採点しても判断に困ることが推測されるようなものであった。非常に多数の受験者を扱う場合には、避けがたい事例であり、採点時における困難が容易に予想される。

記述式解答が受験者の資質をより良く判定できると思われる問題があることは事実であるが、具体的に作題し且つ採点してみると、採点時の困難まで考慮に入れた記述式問題の出題の難しさが改めて実感された。一方、記述式で測ろうとする能力が実は作題の工夫に依り選択式の出題でもかなり測れるのではないかということを感じたのも事実であり、今後の課題である。

また、新しい試験では、段階分けによる成績報告、解答状況（どの問題に解答したか、または解答しなかったか）や答案内容の分類などの情報を利用するなど、評価や報告の仕方が従来と大きく異なる可能性についても考えていく必要がある。

5.5 試作問題の特徴を理解するための研究

立脇洋介

5.5.1 スーパームーン問題

平成27年12月22日に開催された「高大接続システム改革会議」(第9回)において示された「問題イメージ」のうち、数学の問題(以後、「スーパームーン問題」と呼ぶ)をモニター受験者に示し、解答を求めた。

スーパームーン問題は、第2回モニター調査(1月17日、24日)に実施された。この問題は、三角法に関連する3つの問から構成されている。1月17日には、260名を対象として解答時間の上限を25分と設定した。また、1月24日には、67名を対象として解答時間の上限を30分に設定した。

フィージビリティ検証事業における数学作業グループにて、採点案を作成し(公開された解答案から若干の変更を伴っている)、これに基づいて数学の教育経験のある外部協力者に、採点作業を依頼した。

①採点結果

採点案(別冊に収録)に基づき、外部の協力者3名が採点を行った。問題ごとの得点の基準ならびに3名の採点結果による得点率(満点に対する平均点の比率)の平均を表1に示す。問1～問3の3問で構成されていたが、問1と問2は得点率が10%前後と難易度が非常に高かった。ただし、問1は「単位と式という2つの採点基準のうち、式のみ正解」という人が半数ほどであった。問3は得点率が30%程度と他の間に比べると易しかった。

ただし、本調査の参加者は、本年度のセンター試験の数学Ⅰ・数学Aと数学Ⅱ・数学Bも受験し、実際のセンター試験の受験者より平均点が20点前後高い集団である。そのような集団で得点率が10～30%の問題は、共通テストとしては難易度が高すぎたと考えられる。実際の受験生を想定した難易度の設定になるよう、調整していく必要がある。

表1 問題ごとの得点の基準ならびに得点率の平均

	得点の基準	得点率の平均
問1得点	正解を2点, 不正解を0点とした	9.2%
問1式	正解を1点, 不正解を0点とした	53.4%
問1単位	正解を1点, 不正解を0点とした	9.2%
問2得点	平行性, 対称性, 大小関係の合計(0～6点)	15.3%
問2平行性	満点を2点, 部分点を1点, 誤りを0点とした	15.3%
問2対称性	満点を2点, 部分点を1点, 誤りを0点とした	21.1%
問2大小関係	満点を2点, 部分点を1点, 誤りを0点とした	9.7%
問3得点	正解を2点, 不正解を0点とした	31.0%
合計得点	問1～問3の合計得点(0～10点)	17.2%

②信頼性

採点結果の信頼性を検討するために、まず採点者間の相関係数を算出した。3名分の採点結果の類似度である級内相関に加えて、3名のうちの2名のペア相関も算出した。数値が大きいほど、関連が強いことを示す。意味のある関連が見られた箇所にはアスタリスクを付した。結果を表2に示す。

問1は、得点に関する採点結果が完全に一致しており、式や単位でも相関係数が.950前後であった。問3も相関係数が.930以上であった。採点指示を見ると、問1と問3の正解は、「AあるいはB」のように比較的単純かつ明確であるため、採点者によるぶれが少なかったと推測される。

問2は、問1や問3に比べて、相関係数が低い。特に、大小関係は級内相関が.148と非常に低かった。その原因を明らかにするために、採点者別に得点率（満点に対する平均点の比率）を算出した。表3から、採点者aは問2の大小関係の得点率が0.2%と非常に低く、採点が厳しかった。その一方で、採点者cは得点率が20.8%であった。採点指示を見ると、問1や問3に比べて複雑であり、採点者によって基準にばらつきが生じやすかった。また採点者から提出されたレポートにおいて、「想定していた解答方法以外についても基準を設定すべき」という指摘もなされていた（別冊参照）。

表2 採点者間の採点結果の類似度（N=327の相関係数）

	級内相関	ペア相関		
		採点者ab	採点者ac	採点者bc
問1得点	1.000**	1.000**	1.000**	1.000**
問1式	.955**	.969**	.933**	.963**
問1単位	.976**	.982**	.964**	.982**
問2得点	.546**	.720**	.627**	.794**
問2平行性	.702**	.672**	.703**	.803**
問2対称性	.448**	.508**	.322**	.695**
問2大小関係	.148**	.117*	.049	.532**
問3得点	.938**	.930**	.944**	.944**
合計得点	.795**	.876**	.847**	.905**

** $p < 0.01$ * $p < 0.05$

表3 採点者別に見た得点率

	得点率		
	採点者a	採点者b	採点者c
問1得点	9.2%	9.2%	9.2%
問1式	54.4%	53.5%	52.3%
問1単位	8.9%	9.2%	9.5%
問2得点	7.9%	15.6%	22.5%
問2平行性	12.5%	14.7%	18.7%
問2対称性	11.2%	24.0%	28.1%
問2大小関係	0.2%	8.3%	20.8%
問3得点	31.8%	29.4%	31.8%
合計得点	12.9%	17.1%	21.7%

次に問題間の採点結果の類似度（内的一貫性）を検討するために、得点間の相関係数を算出した（表 4）。全般的に中程度以上の相関関係が見られ、共通した能力を測定していると考えられる。ただし、問 1 の採点結果と、問 2 と問 3 の採点結果との間の相関係数は、.200 未満と小さかった。

表 4 問題間の採点結果の類似度（N=327 の相関係数）

	問 1 得 点	問 1 式	問 1 単 位	問 2 得 点	問 2 平 行 性	問 2 対 称 性	問 2 大 小 関 係	問 3 得 点	合 計 得 点
問1得点	1.000**	.301**	.996**	.120*	.070	.126*	.095	.118*	.433**
問1式	.301**	1.000**	.301**	.169**	.098	.205**	.077	.134*	.255**
問1単位	.996**	.301**	1.000**	.127*	.075	.129*	.104	.127*	.440**
問2得点	.120*	.169**	.127*	1.000**	.773**	.829**	.756**	.464**	.841**
問2平行性	.070	.098	.075	.773**	1.000**	.361**	.386**	.272**	.601**
問2対称性	.126*	.205**	.129*	.829**	.361**	1.000**	.570**	.430**	.726**
問2大小関係	.095	.077	.104	.756**	.386**	.570**	1.000**	.413**	.667**
問3得点	.118*	.134*	.127*	.464**	.272**	.430**	.413**	1.000**	.787**
合計得点	.433**	.255**	.440**	.841**	.601**	.726**	.667**	.787**	1.000**

** $p < 0.01$ * $p < 0.05$

③採点時間

採点作業にかかった時間を採点者別にまとめた結果を表 5 に示す。327 名分の時間に加えて 1 名当たりの時間も算出した。1 名あたりおよそ 2~3 分程度かかっていた。

表 5 採点者別の採点時間

	採点時間(分) 1名あたり(秒)	
採点者a	840	154
採点者b	690	127
採点者c	950	174

④解答時間の影響

1 月 17 日の解答中の受験生の様子を観察し、解答時間が不足している可能性が考えられた。そこで、24 日の実験では解答時間の制限を 25 分から 30 分に延長した。解答時間の延長によって得点が異なるかを検討した結果を表 6 に示す。分析の結果、いずれの問題においても、解答時間による差は見られなかった。

表 6 解答時間別に見た得点率

	得点率		統計量	
	25分 (N=253)	30分 (N=66)	<i>t</i>	<i>df</i>
問1得点	9.2%	10.7%	0.37	317
問1式	52.7%	57.0%	0.66	317
問1単位	9.0%	10.7%	0.38	317
問2得点	14.9%	16.9%	0.78	317
問2平行性	14.3%	17.2%	0.77	317
問2対称性	21.7%	20.5%	0.31	317
問2大小関係	9.0%	13.3%	1.83	90.06
問3得点	31.3%	30.3%	0.17	317
合計得点	17.1%	18.3%	0.49	317

5.5.2 試作問題とセンター試験の問題の特性比較

国語と数学の2教科に関して、第2回モニター調査で用いた試作問題とセンター試験の問題の特性を比較することを目的に相関分析を実施した。分析に用いたテストの一覧を表7に示す。表7のうち、思考力に関するテストは、第2回モニター調査の受験者の一部が参加した、第3回モニター調査で実施したものである。

表7 問題の特性比較に使用したテスト一覧

テスト名	説明
試作国語(J3冊子)	
第1問問1	0点, 1点, 2点の3段階。
第1問問2	0点, 3点, 5点, 8点, 10点の5段階。
第2問	0点, 1点, 2点の3段階。
第3問内容	0点, 1点, 2点, 3点, 4点の5段階。
第3問表現	0点, 1点, 2点, 3点, 4点の5段階。
第3問構成	0点, 1点, 2点, 3点, 4点の5段階。
第3問総合	0~12点。
センター国語	
合計	28年度本試と追試の国語全体の合計。400点満点。
論説	28年度本試と追試の論説(第1問)の合計。100点満点。
小説	28年度本試と追試の小説(第2問)の合計。100点満点。
古典	28年度本試と追試の古典(第3問)の合計。100点満点。
漢文	28年度本試と追試の漢文(第4問)の合計。100点満点。
試月合計	スーパームーン問題の合計得点。0~10点。
試作数学マーク(M4冊子)	
合計	75点満点。
第1問	25点満点。
第2問	25点満点。
第3問	25点満点。
試作数学記述(M3冊子)	
合計	75点満点。
第1問	25点満点。
第2問	25点満点。
第3問	25点満点。
センター数学	
合計	28年度本試と追試の数学全体の合計。400点満点。
数1	28年度本試と追試の数学1Aの合計。200点満点。
数2	28年度本試と追試の数学2Bの合計。200点満点。
センター英語	
英合計	28年度本試と追試の英語全体の合計。500点満点。
英筆記	28年度本試と追試の英語筆記の合計。400点満点。
英リ	28年度本試と追試の英語リスニングの合計。100点満点。
思考力	
批判的思考力テスト	久原・井上・波多野(1983)。80点満点。
論理的判断力	日弁連法務研究財団の法科大学院全国統一適性試験(2015年第1回)。24点満点。
分析的判断力	日弁連法務研究財団の法科大学院全国統一適性試験(2015年第1回)。24点満点。
長文読解力	日弁連法務研究財団の法科大学院全国統一適性試験(2015年第1回)。24点満点。
論理的な思考	国立教育政策研究所による平成24年特定の課題に関する調査。記号などで解答する10問の正解数。10点満点。

①国語

国語の結果を表 8 と表 9 に示す。数値の絶対値が大きいほど、関連が強いことを示す。

全体的な傾向として試作問題の国語は、「試作問題国語の他の問」との関連が弱かった。

「センター試験の国語」との間でも相関係数が.20 未満と低く、弱い関連しか見られなかった。さらに、センター試験の数学や英語との間でも、試作問題国語はセンター試験国語に比べて関連が弱かった（試作問題とセンター試験英語との相関は.10～.20）。

思考力を測定するテストとの関連を見ると、センター試験は比較的「長文読解力」に近い能力を測定していると考えられる。一方、試作問題第 1 問の問 1（交通事故のグラフの問題）は、「長文読解力」に加えて「論理的な思考」との間に関連が見られた。また、試作問題第 3 問は、「批判的思考力」との間に関連が見られた。

ただし、試作問題第 1 問の問 2 と試作問題第 2 問（連動型複数選択問題）については、センター試験の国語の一部や英語と弱い関連が見られたものの、思考力テストとは関連が見られなかった。

以上の結果から、試作問題国語は新しい能力を測定しているのか、それとも測定や採点上の課題が残されているのか、慎重に検討を進める必要がある。

②数学

数学の結果を表 10 と表 11 に示す。全体的な傾向として、スーパームーン問題、マーク式、記述式のいずれも試作問題の数学は、センター試験の数学と強い関連を示していた。

「試作問題数学の他の問」との関連を見ると、マーク式、記述式ともに中程度の相関を示していた。つまり各問において多少異なる能力を測定しているものの、全体として共通する数理的な能力を測定していると考えられる。

センター試験の国語や英語との関連を見ると、試作問題はセンター試験の数学と類似した傾向を示し、国語とは弱い相関を示し、英語とは中程度の相関を示していた。思考力を測定するテストとの関連を見ると、試作問題の数学もセンター試験の数学も共通して、「批判的思考力」「論理的な思考」「論理的判断力」「分析的判断力」との間に中程度の相関を示した。

以上から他のテストとの関連に関して、試作問題とセンター試験で結果が類似していた。このような結果となった理由としては、数学的な能力が比較的安定しており、解答方法や出題内容の影響をそれほど受けなかったためと考えられる。

表8 試作問題とセンター試験の問題との特性比較 (国語とセンター試験の他教科との相関係数)

	試作国語					センター国語					N		
	第1問問1	第1問問2	第2問	第3問内容	第3問表現	第3問構成	第3問総合	合計	論説	小説		古典	漢文
試国第1問問1	1.000	.032	-.014	-.014	.056	.010	.011	.112*	.114*	.015	.103*	.097	363
試国第1問問2	.032	1.000	-.036	.009	.027	.009	.012	.102	.011	-.031	.178**	.098	363
試国第2問	-.014	-.036	1.000	.001	-.054	-.031	-.031	.099	.103*	.092	.062	.061	363
試国第3問内容	-.014	.009	.001	1.000	.803**	.903**	.941**	.115*	.012	.057	.140**	.114*	363
試国第3問表現	.056	.027	-.054	.803**	1.000	.887**	.933**	.185**	.090	.064	.198**	.180**	363
試国第3問構成	.010	.009	-.031	.903**	.887**	1.000	.979**	.131*	.038	.057	.148**	.129*	363
試国第3問総合	.011	.012	-.031	.941**	.933**	.979**	1.000	.139**	.042	.055	.157**	.141**	363
セ国合計	.112*	.102	.099	.115*	.185**	.131*	.139**	1.000	.738**	.628**	.840**	.831**	363
セ国論説	.114*	.011	.103*	.012	.090	.038	.042	.738**	1.000	.410**	.435**	.499**	363
セ国小説	.015	-.031	.092	.057	.064	.057	.055	.628**	.410**	1.000	.370**	.315**	363
セ国古典	.103*	.178**	.062	.140**	.198**	.148**	.157**	.435**	.410**	.370**	1.000	.627**	363
セ国漢文	.097	.098	.061	.114*	.180**	.129*	.141**	.499**	.435**	.370**	.627**	1.000	363
セ国合計	.047	.073	.066	.032	.068	.044	.054	.292**	.203**	.043	.258**	.336**	363
セ数1	.074	.098	.075	.033	.082	.050	.059	.332**	.237**	.061	.290**	.372**	363
セ数2	.024	.050	.054	.030	.052	.036	.046	.242**	.163**	.027	.215**	.285**	363
セ英合計	.147**	.155**	.104*	.140**	.193**	.157**	.165**	.660**	.495**	.275**	.552**	.641**	363
セ英筆記	.140**	.150**	.093	.138**	.189**	.159**	.164**	.651**	.502**	.260**	.541**	.633**	363
セ英リ	.153**	.150**	.132*	.124*	.177**	.119*	.139**	.584**	.378**	.290**	.501**	.566**	363
試月合計得点	.044	.021	.102	.038	.049	.022	.041	.223**	.176**	.069	.208**	.204**	319
試数又第1問	.044	.045	.051	.024	.089	.049	.051	.157*	.139	-.019	.172*	.136	183
試数又第2問	-.012	.033	.002	.005	.078	.055	.044	.204**	.151*	.076	.194**	.165*	183
試数又第3問	.083	.020	.091	.071	.156*	.094	.105	.148*	.100	.011	.173*	.122	183
試数又第2問	.019	.056	.017	-.026	-.023	-.028	-.030	.048	.097	-.108	.061	.057	183
試数記第1問	.044	.072	.027	.104	.113	.130	.128	.291**	.196**	-.028	.259**	.401**	180
試数記第2問	.077	.092	.033	.136	.142	.147*	.149*	.252**	.153*	-.023	.268**	.316**	180
試数記第3問	-.039	.054	.042	.012	.055	.044	.042	.147*	.087	-.049	.131	.235**	180
試数記第3問	.072	.042	-.002	.113	.092	.131	.128	.302**	.224**	.000	.241**	.410**	180

** p<0.01 * p<0.05

表9 試作問題とセンター試験の問題との特性比較 (国語と思考力テストとの順位相関係数)

	試作国語					センター国語					N		
	第1問問1	第1問問2	第2問	第3問内容	第3問表現	第3問構成	第3問総合	合計	論説	小説		古典	漢文
批判的思考力	.240	-.065	-.092	.236	.250*	.253*	.249*	.174	.062	.000	.270*	.191	64
論理的な思考	.272*	-.074	-.112	.064	.102	.076	.110	.163	.161	.107	.242	.012	64
論理的判断力	.286*	.049	-.183	.124	.021	.091	.061	.240	.161	.101	.278*	.148	62
分析的判断力	.210	.040	-.084	.033	.061	.095	.061	.130	.013	.093	.056	.237	62
長文読解力	.262*	.108	.072	.150	.105	.138	.135	.324*	.286*	.233	.307*	.202	62

** p<0.01 * p<0.05

表10 試作問題とセンター試験の問題との特性比較(数学とセンター試験の他教科との相関係数)

	試作問題とセンター試験の問題との特性比較(数学とセンター試験の他教科との相関係数)			試作問題とセンター試験の問題との特性比較(数学とセンター試験の他教科との相関係数)			試作問題とセンター試験の問題との特性比較(数学とセンター試験の他教科との相関係数)					
	試作月 (N=319)	合計	第1問	第2問	第3問	合計	第1問	第2問	第3問	合計	数1	数2
試月合計得点	1.000	.432**	.239**	.443**	.331**	.522**	.321**	.519**	.387**	.605**	.511**	.599**
試数で合計	.432**	1.000	.721**	.809**	.877**	-	-	-	-	.794**	.786**	.749**
試数で第1問	.239**	.721**	1.000	.390**	.490**	-	-	-	-	.547**	.575**	.551**
試数で第2問	.443**	.809**	.390**	1.000	.543**	-	-	-	-	.624**	.643**	.599**
試数で第3問	.331**	.877**	.490**	.543**	1.000	-	-	-	-	.655**	.799**	.749**
試数記合計	.522**	-	-	.543**	1.000	1.000	.672**	.851**	.847**	.803**	.519**	.726**
試数記第1問	.321**	-	-	.390**	.490**	.672**	1.000	.373**	.392**	.519**	.647**	.580**
試数記第2問	.519**	-	-	.390**	.490**	.373**	.373**	1.000	.572**	.572**	.726**	.647**
試数記第3問	.387**	-	-	.390**	.490**	.392**	.572**	.572**	1.000	.649**	.580**	.644**
七数合計	.605**	.794**	.547**	.660**	.698**	.803**	.519**	.726**	.649**	1.000	.938**	.969**
七数1	.551**	.786**	.575**	.624**	.694**	.714**	.457**	.647**	.580**	.938**	1.000	.825**
七数2	.599**	.749**	.493**	.643**	.655**	.799**	.521**	.723**	.644**	.969**	.825**	1.000
七国合計	.176**	.196**	.153*	.087	.224**	.139	.151*	.100	.097	.203**	.237**	.163**
七国論説	.069	-.028	-.023	-.049	.000	-.019	.076	.011	-.108	.043	.061	.027
七国小説	.208**	.259**	.268**	.131	.241**	.172*	.194**	.173*	.061	.258**	.290**	.215**
七国古典	.204**	.401**	.316**	.235**	.410**	.136	.165*	.122	.057	.336**	.372**	.285**
七国漢文	.223**	.291**	.252**	.147*	.302**	.157*	.204**	.148*	.048	.292**	.332**	.242**
七英合計	.336**	.362**	.334**	.202**	.347**	.414**	.320**	.348**	.325**	.484**	.508**	.431**
七英筆記	.336**	.371**	.338**	.210**	.356**	.428**	.333**	.355**	.340**	.497**	.519**	.445**
七英リ	.279**	.257**	.259**	.126	.248**	.283**	.208**	.262**	.204**	.341**	.370**	.295**
試国第1問問1	.044	.044	.077	-.039	.072	.044	-.012	.083	.019	.047	.074	.024
試国第1問問2	.021	.072	.092	.054	.042	.045	.033	.020	.056	.073	.098	.050
試国第2問	.102	.027	.033	.042	-.002	.051	.002	.091	.017	.066	.075	.054
試国第3問内容	.038	.104	.136	.012	.113	.024	.005	.071	-.026	.032	.033	.030
試国第3問表現	.049	.113	.142	.055	.092	.089	.078	.156*	-.023	.068	.082	.052
試国第3問構成	.022	.130	.147*	.044	.131	.049	.055	.094	-.028	.044	.050	.036
試国第3問総合	.041	.128	.149**	.042	.128	.051	.044	.105	-.030	.054	.059	.046

** p<0.01 * p<0.05

表11 試作問題とセンター試験の問題との特性比較(数学と思考力テストとの順位相関係数)

	試作数学マーク(N=22)			試作数学記述(N=42)			センター数学(N=64)					
	合計	第1問	第2問	合計	第1問	第2問	合計	数1	数2			
批判的思考力	.309*	.426*	.252	.478*	.147	.167	-.073	.108	.206	.269*	.277*	.266*
論理的な思考	.366**	.656**	.408	.606**	.530*	.375*	.289	.237	.366*	.496**	.472**	.487**
論理的判断力	.254	.662**	.418	.565**	.337	.457**	.264	.334*	.463**	.541**	.548**	.504**
分析的判断力	.335*	.303	.175	.214	.114	.370*	.207	.292	.314*	.541**	.506**	.541**
長文読解力	.021	.460*	.697**	.305	.219	.125	-.048	.123	.177	.197	.201	.177

** p<0.01 * p<0.05

5.5.3 アイトラッカーを用いた注視点追跡実験

国語の能力によって問題文の読み方に違いがあるかを明らかにするために、第3回モニター調査でアイトラッカーを用いて注視点の追跡実験を行った。

①実験参加者

国語の能力の高い人としては、「センター試験国語の本試と追試の合計点が340点（400点満点）以上」かつ「試作問題国語の第1問の問1が満点」だった人から3名を選出した（High群）。一方、国語の能力の低い人としては、「センター試験国語の本試と追試の合計点が280点（400点満点）未満」かつ「試作問題国語の第1問の問1が×」だった人から3名を選出した（Low群）。

②手続き

実験の説明、調整、練習の後、パソコンの画面上に縦書きで2ページ分の文章を提示した（図3と図4を参照）。画面の右側は素材文で、左側に解答する問題を提示した。答えが分かった際には、口頭で解答を求めた。この間、協力者の目に弱い赤外線を当て、視線を記録した（図1を参照）。計測開始から終了までの間の視線の位置が記録され、図2のように再現することができる。

全ての協力者に2種類の問題を提示した。問題1は、平成27年度のセンター試験国語本試の論説問題の一部である。問題2は、試作問題国語の第1問の問1の一部である。



図1 アイトラッカーを用いた注視点追跡実験の様子

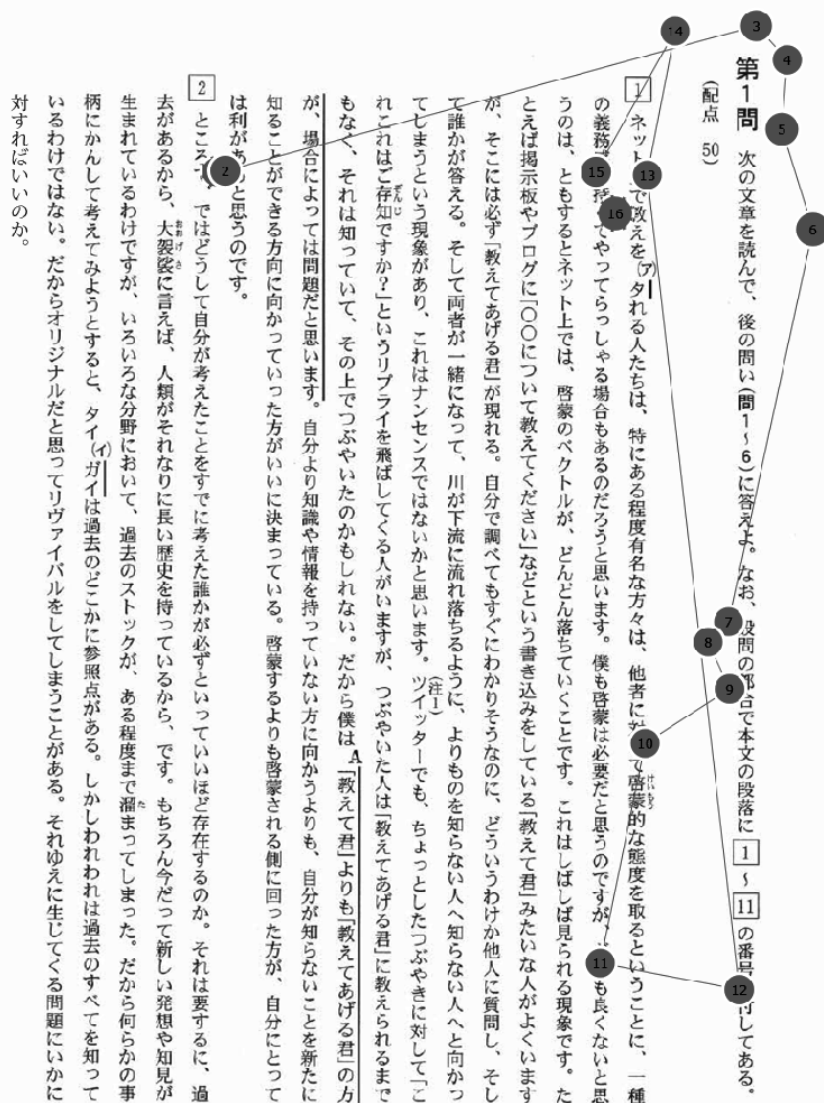


図2 注視点の移動の例

③結果

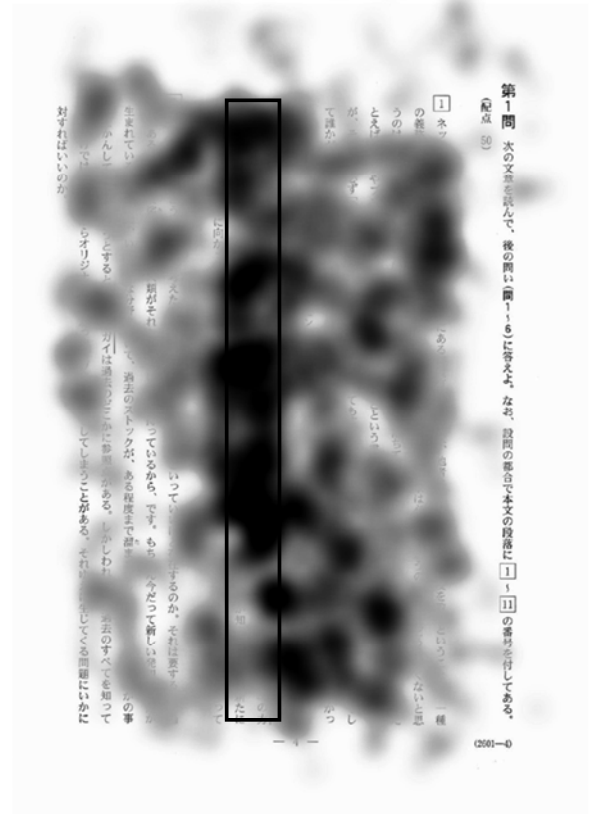
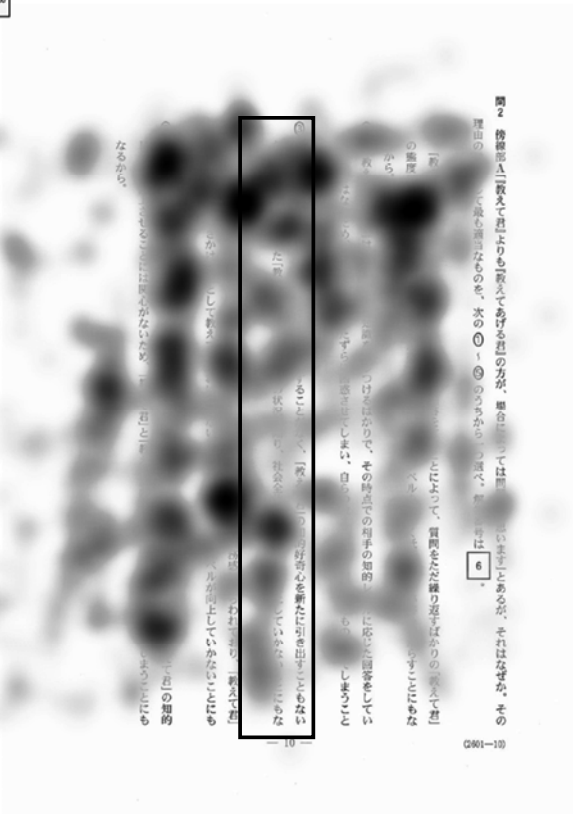
High群とLow群とに分けて、注視した時間(注視量)を図示した。問題1は図5に、問題2は図6に示す。図において濃い色の部分ほど、注視した時間が長いことを示す。また、正解選択肢や正解のために必要な情報は四角で囲んで示してある。

センター試験の問題について、High群は「正解選択肢」を中心に読んでいた。一方Low群は素材文の該当箇所を中心に、何度も読み返していた。試作問題のグラフ部分でも、同様の傾向が見られ、High群は正解に必要なグラフのタイトルのみを見ていたのに対して、Low群全てのグラフのタイトルをまんべんなく見ていた。以上の結果から、国語の能力の高い人は、センター試験のような文章や試作問題のようなグラフの中から必要な情報を選択し、集中的に読み取っていると推測される。

ただし、本実験はHigh群とLow群とも3名しかおらず、能力と視線との関連を探索的に検討したに過ぎない。人数や素材文を拡大するなどし、詳細な検討が必要である。



Low群



High群

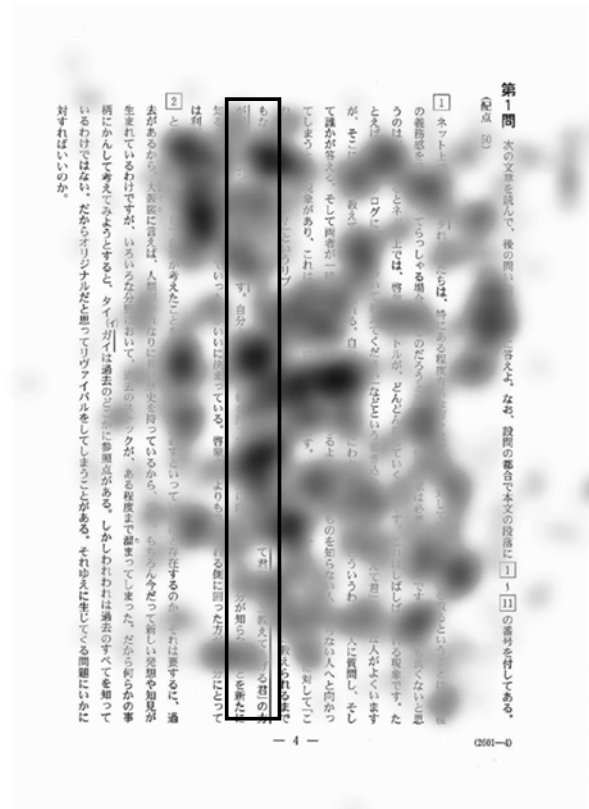
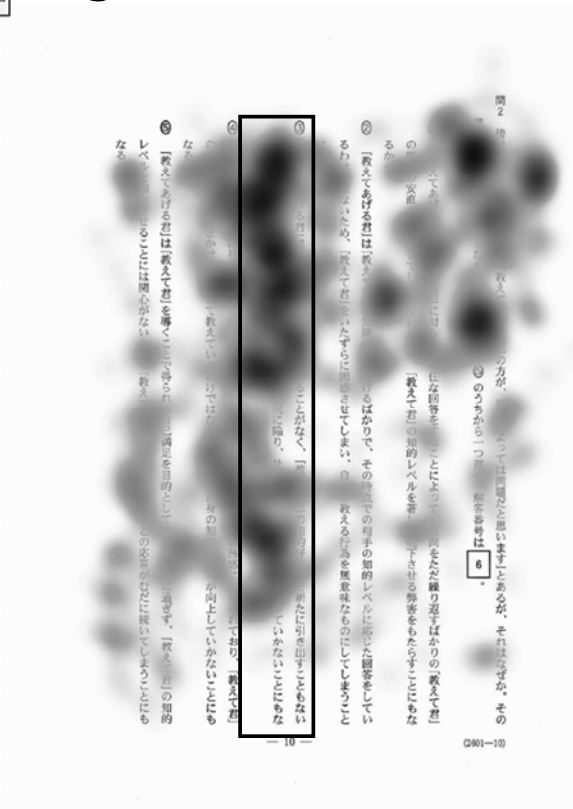


図5 センター試験国語の問題文を読む際の注視量



Low群

問1 Bさんは、下線部(a)つまり以下、どのような内容を述べることになるか。空欄アに出るはまる適切な内容を四十字以内で書け(句読点を含む)。

Aさん「交通のつり合いが、平成二年(一九九〇年)以降減少傾向になっているのは、交通安全に関する国民の意識が向上しているからだと思います。」

その資料の資料として、「交通違反で検挙された人数の推移が分かる資料」があります。その資料を見れば、交通違反の割合が少なくなっていることが分かるはずだと思いませんか。

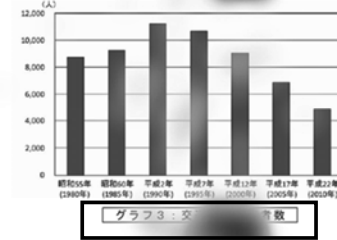
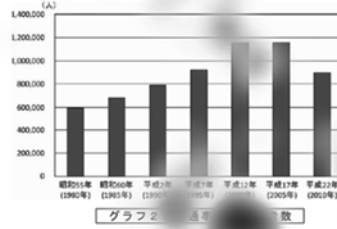
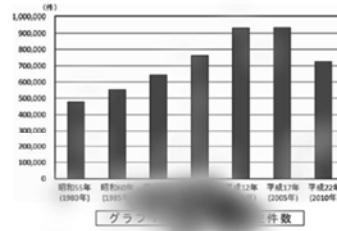
Bさん「私は、この資料から、交通違反の割合が少なくなっていることが分かると思いませんか。」

文は、(a) 交通違反の割合が少なくなっていることが分かると思いませんか。

例えは、(a) 交通違反の割合が少なくなっていることが分かると思いませんか。

空欄ア

空欄ア



第1問 次の文章とグラフを読み、後の問い(問1と2)に答えよ。

次に示すのは、警察庁事故統計資料に基づいて作成された交通事故の発生件数、負傷者数、死者数のグラフと、この三つのグラフを見て、交通事故の死者数が他よりも早く、平成二年(一九九〇年)以降減少傾向になっていることについて、四人の高校生が行った話し合いの一部である。



High群

問1 Bさんは、下線部(a)つまり以下、どのような内容を述べることになるか。空欄アに出るはまる適切な内容を四十字以内で書け(句読点を含む)。

Aさん「交通のつり合いが、平成二年(一九九〇年)以降減少傾向になっているのは、交通安全に関する国民の意識が向上しているからだと思います。」

その資料の資料として、「交通違反で検挙された人数の推移が分かる資料」があります。その資料を見れば、交通違反の割合が少なくなっていることが分かるはずだと思いませんか。

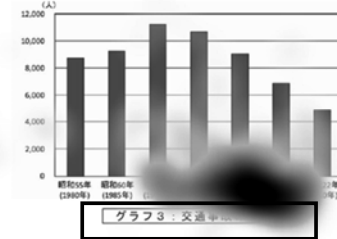
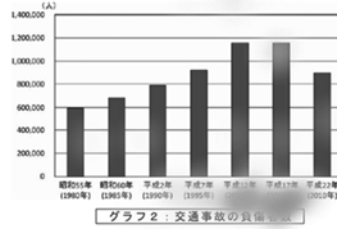
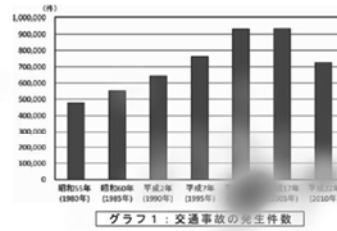
Bさん「私は、この資料から、交通違反の割合が少なくなっていることが分かると思いませんか。」

文は、(a) 交通違反の割合が少なくなっていることが分かると思いませんか。

例えは、(a) 交通違反の割合が少なくなっていることが分かると思いませんか。

空欄ア

空欄ア



第1問 次の文章とグラフを読み、後の問い(問1と2)に答えよ。

次に示すのは、警察庁事故統計資料に基づいて作成された交通事故の発生件数、負傷者数、死者数のグラフと、この三つのグラフを見て、交通事故の死者数が他よりも早く、平成二年(一九九〇年)以降減少傾向になっていることについて、四人の高校生が行った話し合いの一部である。

図6 試作問題国語の問題文を読む際の注視量
85

c 新テスト実施環境の開発

c-1 CBT実施のプロトタイプシステムの開発

1 情報端末を用いたテスト・システムの運用調査

1.1 概要

本研究期間中に四回の CBT による実験調査をおこなった。
実験調査の目的は、

1. 記述解答データの収集
2. 仕様する情報端末の特性と実施運用の関係性の検討
3. テスト実行アプリケーションの検討と改善

の3点である。

まず、それぞれ四回の調査結果について報告する。

なお、本研究期間中に用いた情報端末の特徴については以下のとおりである。なお、OS はすべて windows の enterprise edition であった。第一回～第三回までは windows8.1 を用い、第四回は windows10 を用いている。

表 1.1: 利用した情報端末

製品	画面サイズ	CPU	メモリ	SDD	特徴
A 社	10.1inch(1366* 768)	1.8GHz	2GB	64GB	着脱式キーボード
B 社	10.1inch(1366* 768)	1.3-1.8GHz	2GB	32GB	
C 社	12.5inch(1920*1080)	0.8-2.0GHz	4GB	128GB	種々のプリインストールパッケージあり
D 社	12.0inch(2160*1440)	1.5GHz	4GB	64GB	着脱式キーボード
E 社	13.3inch(1920*1080)	2.2GHz	4GB	256GB	

1.2 第一回試行調査 (20150725)

1.2.1 実施概要

第一回試行調査は、2015 年 7 月 25 日におこなわれた。試験運用形式は CBT と紙筆試験の両方であった。その他の概要については、以下の表に示す。また、試験当日の CBT 実施風景と紙筆試験の実施風景はそれぞれ以下のものであった。

表 1.2: 試行調査の実施概要

項目	
実施日時	2015年7月25日
実施場所	NTT データ駒場
対象者	センター試験受験経験者
試験時間(分)	140
調査完了人数	253
完了人数(CBT)	97
完了人数(紙筆)	156
試験問題数	22



図 1.1: CBT 実施風景



図 1.2: 紙筆試験実施風景

1.2.2 試験問題

実施した試験問題は、下記のとおりであった。解答形式はすべて記述解答形式であった。

表 1.3: 実施した試験問題

#	試験問題	基準時間	字数
1	国語	08:00	20 字以内
2	生物	06:00	30 字以内
3	生物	07:00	40 字以上 60 字以内
4	化学	06:00	30 字以上 50 字以内
5	地理	10:00	単語・20 文字以内
6	国語	08:00	20 字以上 40 字以内
7	化学	06:00	30 字以内
8	化学	05:00	15 字以内・5 字以内
9	日本史	05:00	15 字以内・15 字以内
10	世界史	06:00	20 字以上 40 字以内
	休憩	10:00	
11	世界史	05:00	30 字以内
12	生物	06:00	30 字以内
13	世界史	06:00	20 字以上 40 字以内
14	化学	05:00	10 字以内
15	国語	12:00	30 字以内・20 字以上 40 字以内
16	生物	10:00	30 字以上 50 字以内・20 字以内
17	物理	09:00	式を入力・式を入力
18	日本史	07:00	30 字以内
19	化学	05:00	10 字以内
20	数学	04:00	式を入力
21	日本史	06:00	20 字以上 30 字以内
22	世界史	05:00	30 字以上 50 字以内

1.2.3 受験者集団

受験者集団は下記のとおりであった。

表 1.4: 受験者集団の出身(在籍)大学

大学名	人数	割合	大学名	人数	割合	大学名	人数	割合
早稲田大学	26	10.3%	東京外国語大学	6	2.4%	筑波大学	2	0.8%
慶應義塾大学	26	10.3%	東京工業大学	5	2.0%	津田塾大学	2	0.8%
上智大学	26	10.3%	お茶の水女子大学	4	1.6%	専修大学	2	0.8%
東京大学	25	9.9%	電気通信大学	4	1.6%	大妻女子大学	2	0.8%
法政大学	11	4.3%	学習院大学	3	1.2%	國學院大學	2	0.8%
立教大学	10	4.0%	東京農業大学	3	1.2%	ルーテル学院大学	2	0.8%
中央大学	9	3.6%	成城大学	3	1.2%	東京農工大学	2	0.8%
青山学院大学	8	3.2%	千葉大学	2	0.8%	神奈川大学	2	0.8%
一橋大学	7	2.8%	埼玉大学	2	0.8%	横浜国立大学	1	0.4%
日本大学	7	2.8%	首都大学東京	2	0.8%	その他	47	18.6%

1.2.4 CBT 運用結果

CBT 運用時におこったトラブルは下記のとおりであった。

表 1.5: CBT 運用時のトラブル内容

項目	値	原因
実施前トラブル	1	着脱式キーボードに起因
実施中トラブル	3	不明
試験が完了できなかったもの	3	

1. 今回、CBT において試験完了できなかったものが 3 件。実施前トラブルが 1 件。
2. A 社製端末において、日本史の解答中に初期画面に戻る事故が 2 件。再現性なし。また、生物の問題に回答中に同様の事故が 1 件発生。これも再現性なし。
3. 実施前トラブルは B 社製の着脱式キーボードの不具合であった。
4. 事故発生後は、紙筆試験に切り替えて調査を完了した。なお、解答済みデータについては、3 つのタブレットともに保存されていた。
5. データ消失は回答中の試験問題に対してのみ。
6. なお、これまでの CBT トラブルは、試験実施中については計 7 回の実施 (366 名) で試験中のトラブルが 0 件。実施前のトラブル発見が 4 件であった。
7. これまでの総累積では、試験実施中のトラブルが 8 回実施 (619 名) 中で 3 件。実施前のトラブルが 5 件である。(A 社 3 件 + 4 件, B 社 1 件)

1.3 第二回試行調査 (20151031)

第二回試行調査は、2015 年 10 月 31 日におこなわれた。試験運用形式は CBT と紙筆試験の両方であった。その他の概要については、以下の表に示す。また、試験当日の CBT 実施風景は以下のものであった。

表 1.6: 試行調査の実施概要

項目	
実施日時	2015年10月31日
対象者	国立大学1年生
試験時間(分)	85
調査完了人数	293
完了人数(CBT)	88
完了人数(紙筆)	205
試験問題数	20



図 1.3: CBT 実施風景



図 1.4: CBT 実施風景

1.3.1 試験問題

実施した試験問題は、下記のとおりであった。解答形式はすべて記述解答形式であった。

表 1.7: 実施した試験問題

#	試験問題	基準時間	字数
1	国語 02	08:00	20 字以内
2	生物 01	06:00	30 字以内
3	生物 02	07:00	40 字以上 60 字以内
4	化学 03	06:00	30 字以上 50 字以内
5	地理 01A,B	10:00	単語・20 文字以内
6	国語 01	08:00	20 字以上 40 字以内
7	化学 05	06:00	30 字以内
8	化学 01A,B	05:00	15 字以内・5 字以内
9	日本史 01A,B	05:00	15 文字以内・15 文字以内
10	世界史 01	05:00	30 文字以内
11	生物 04	06:00	30 字以内
12	化学 02	05:00	10 字以内
13	国語 03A,B	12:00	30 字以内・20 字以上 40 字以内
14	生物 03A,B	10:00	30 字以上 50 字以内・20 字以内
15	日本史 02	07:00	30 字以内
16	化学 04	05:00	10 字以内
17	世界史 02	05:00	30 字以上 50 字以内
18	日本史 03	06:00	20 字以上 30 字以内
19	世界史 03	06:00	20 字以上 40 字以内
20	世界史 04	06:00	20 字以上 40 字以内

1.3.2 CBT 運用結果

CBT 運用時におこったトラブルは下記のとおりであった。

表 1.8: CBT 運用時のトラブル内容

項目	値	原因
実施前トラブル	0	
実施中トラブル	1	試験実施アプリケーションが強制的に終了 (A 社製)
試験が完了できなかったもの	3	

1. CBT において試験完了できなかったものが 1 件。実施前トラブルが 0 件。
2. A 社製端末において、日本史の解答中に初期画面に戻る事故が 1 件。再現性なし。
3. 実施前トラブルはなし。
4. 事故発生後は、紙筆試験に切り替えて調査を完了した。なお、解答済みデータについては、タブレットに保存されていた。
5. データ消失は回答中の試験問題に対してのみ。

6. これまでの CBT 総累積では，試験実施中のトラブルが9回実施(708名)中で4件．実施前のトラブルが5件である．A社4件(実施前4件)，B社0件(実施前1件)である．

1.4 第三回試行調査(20151114)

第三回試行調査は，2015年11月14日におこなわれた．試験運用形式はCBTと紙筆試験の両方であった．その他の概要については，以下の表に示す．また，試験当日のCBT実施風景は以下のものであった．

表 1.9: 試行調査の実施概要

項目	
実施日時	2016年11月14日
対象者	国立大学1年生
試験時間(分)	85
調査完了人数	191
完了人数(CBT)	88
完了人数(紙筆)	103
試験問題数	20



図 1.5: CBT 実施風景



図 1.6: 紙筆試験実施風景

1.4.1 試験問題

実施した試験問題は，下記のとおりであった．解答形式はすべて記述解答形式であった．

表 1.10: 実施した試験問題

#	試験問題	基準時間	字数
1	国語 02	08:00	20 字以内
2	生物 01	06:00	30 字以内
3	生物 02	07:00	40 字以上 60 字以内
4	化学 03	06:00	30 字以上 50 字以内
5	地理 01A,B	10:00	単語・20 文字以内
6	国語 01	08:00	20 字以上 40 字以内
7	化学 05	06:00	30 字以内
8	化学 01A,B	05:00	15 字以内・5 字以内
9	日本史 01A,B	05:00	15 文字以内・15 文字以内
10	世界史 01	05:00	30 文字以内
11	生物 04	06:00	30 字以内
12	化学 02	05:00	10 字以内
13	国語 03A,B	12:00	30 字以内・20 字以上 40 字以内
14	生物 03A,B	10:00	30 字以上 50 字以内・20 字以内
15	日本史 02	07:00	30 字以内
16	化学 04	05:00	10 字以内
17	世界史 02	05:00	30 字以上 50 字以内
18	日本史 03	06:00	20 字以上 30 字以内
19	世界史 03	06:00	20 字以上 40 字以内
20	世界史 04	06:00	20 字以上 40 字以内

1.4.2 CBT 運用結果

CBT 運用時におこったトラブルは下記のとおりであった。

表 1.11: CBT 運用時のトラブル内容

項目	値	原因
実施前トラブル	1	着脱式キーボードに起因 (A 社製)
実施中トラブル	3	試験実施アプリケーションが強制的に終了 (A 社製)

1. CBT において試験完了できなかったものが 3 件。実施前トラブルが 1 件。
2. A 社製端末において、解答中に初期画面に戻る事故が 3 件。再現性なし。試験問題については日本史が 1 件、後は生物が 2 件。
3. 実施前トラブルは A 社製の着脱式キーボードの不具合であった。
4. 事故発生後は、紙筆試験に切り替えて調査を完了した。なお、解答済みデータについては、3 つのタブレットともに保存されていた。
5. データ消失は回答中の試験問題に対してのみ。

6. これまでの CBT 総累積では，試験実施中のトラブルが 12 回実施 (796 名) 中で 7 件，実施前のトラブルが 6 件である．A 社 7 件 (実施前 5 件)，B 社 0 件 (実施前 1 件) である．



図 1.7: CBT 実施風景



図 1.8: CBT 実施風景

1.5 第四回試行調査 (20160117/20160124)

1.5.1 実施概要

第四回試行調査は，2016 年 1 月 17 日と 1 月 24 日におこなわれた．試験運用形式は CBT と紙筆試験の両方であった．その他の概要については，以下の表に示す．また，試験当日の CBT 実施風景と紙筆試験の実施風景をそれぞれ以下のものであった．

表 1.12: 試行調査の実施概要

項目	
実施日時	2016 年 1 月 17 日と 1 月 24 日
実施場所	NTT データ駒場研修所
対象者	国立大学 1 年生
試験時間 (分)	55
調査完了人数	343
完了人数 (CBT)	135
完了人数 (紙筆)	208
試験問題数	20



図 1.9: CBT 実施風景



図 1.10: 紙筆試験実施風景

1.5.2 データ回収方法

また、第四回の調査においては解答データ収集システムを新たに構築し、その有用性について確認をおこなった。

これまでは、解答データはタブレット端末に入れた状態でコア・センターまで運搬した上で解答データを吸い上げるという方法が取られていた。しかし、それでは運搬中に解答データ消失のリスクがある。そこで、CBTによる試験が終了後、解答データをQRコードへ変換し、その画像データを送信するという方式が採用された。画像データは図のようなルーターで保存された。併せて、画面にQRコードを表示するという方式も併用された。図のように画面に表示することによって、別途QRコードリーダーを介してデータを収集することも可能になる。

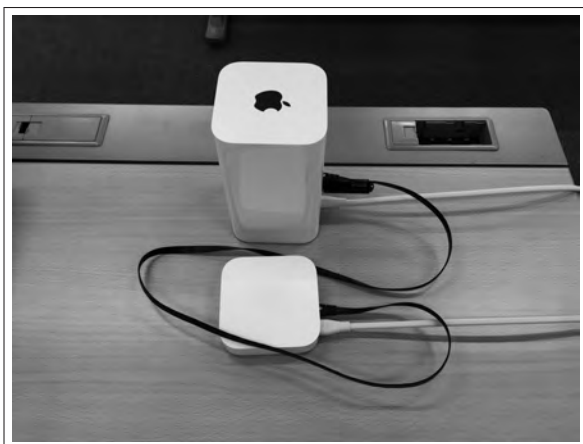


図 1.11: 実施場所に配置したルーター



図 1.12: 終了画面に表示される解答 QR コード

1.5.3 試験問題

実施した試験問題は、下記のとおりであった。解答形式はすべて記述解答形式であった。

表 1.13: 実施した試験問題

#	試験問題	基準時間	字数
1	国語 02	08:00	20 字以内
2	生物 01	06:00	30 字以内
3	生物 02	07:00	40 字以上 60 字以内
4	化学 03	06:00	30 字以上 50 字以内
5	地理 01A,B	10:00	単語・20 文字以内
6	国語 01	08:00	20 字以上 40 字以内
7	化学 05	06:00	30 字以内
8	化学 01A,B	05:00	15 字以内・5 字以内
9	日本史 01A,B	05:00	15 文字以内・15 文字以内
10	世界史 01	05:00	30 文字以内
11	生物 04	06:00	30 字以内
12	化学 02	05:00	10 字以内
13	国語 03A,B	12:00	30 字以内・20 字以上 40 字以内
14	生物 03A,B	10:00	30 字以上 50 字以内・20 字以内
15	日本史 02	07:00	30 字以内
16	化学 04	05:00	10 字以内
17	世界史 02	05:00	30 字以上 50 字以内
18	日本史 03	06:00	20 字以上 30 字以内
19	世界史 03	06:00	20 字以上 40 字以内
20	世界史 04	06:00	20 字以上 40 字以内
21	アンケート		

1.5.4 CBT 運用結果

CBT 運用時におこったトラブルは下記のとおりであった。

表 1.14: CBT 運用時のトラブル内容

項目	値	原因
実施前トラブル	1	着脱式キーボードに起因
実施中トラブル	3	試験実施アプリケーションが強制的に終了

1. CBT において試験完了できなかったものが 3 件。実施前トラブルが 1 件。
2. E 社製端末において、解答中に初期画面に戻る事故が 3 件。再現性なし。
3. 実施前トラブルは D 社製の着脱式キーボードの不具合であった。
4. 事故発生後は、紙筆試験に切り替えて調査を完了した。なお、解答済みデータについては、3 つのタブレットともに保存されていた。

5. これまでの CBT 総累積では、試験実施中のトラブルが 13 回実施 (931 名) 中で 10 件。実施前のトラブルが 7 件である。A 社 7 件 (実施前 5 件), B 社 0 件 (実施前 1 件), D 社 (実施前 1 件), E 社 3 件である。

1.5.5 CBT 実施後のアンケート

アンケート実施概要

試行調査の第四回時に実施した CBT による試験に関するアンケート結果について以下に示す。それぞれ、使用したタブレットの画面サイズによってアンケート結果を別けている。なお、10inch サイズの使用者のサンプルサイズは 88 台であった。また、12-13inch サイズの使用者が 56 台であった。そのうち、19 名は紙筆試験と CBT の両方を受験した者であった。

質問項目は、

1. CBT において 1 問ごとの制限時間設定は解答に影響がありましたか？
2. CBT においてキーボード入力によって解答に悪影響はあったと思いますか？
3. CBT において画面の見にくさによって解答に悪影響はあったと思いますか？
4. CBT において操作のしにくさによって解答に悪影響はあったと思いますか？

の 4 つの項目と、それぞれに対する理由を書かせるものであった。

さらに、紙筆試験と CBT の両方受けたものに対しては、「試験問題が読みやすいのはどちらか?」「解答しやすいのはどちらか?」の 2 つのアンケートを追加で実施した。

アンケート結果

アンケートの結果を以下に示す。



図 1.13: Q. CBT において 1 問ごとでの制限時間設定は解答に影響がありましたか？

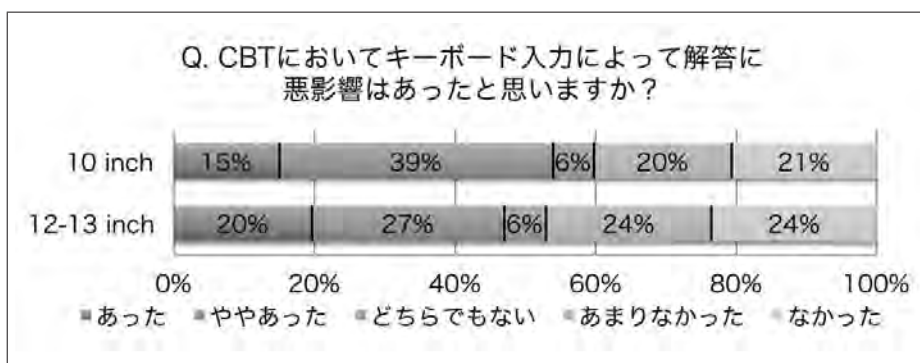


図 1.14: Q. CBT においてキーボード入力によって解答に悪影響はあったと思いますか？

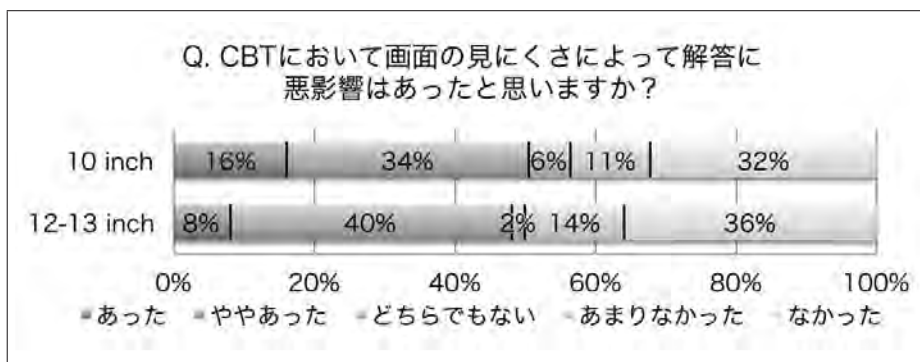


図 1.15: Q. CBT において画面の見にくさによって解答に悪影響はあったと思いますか？

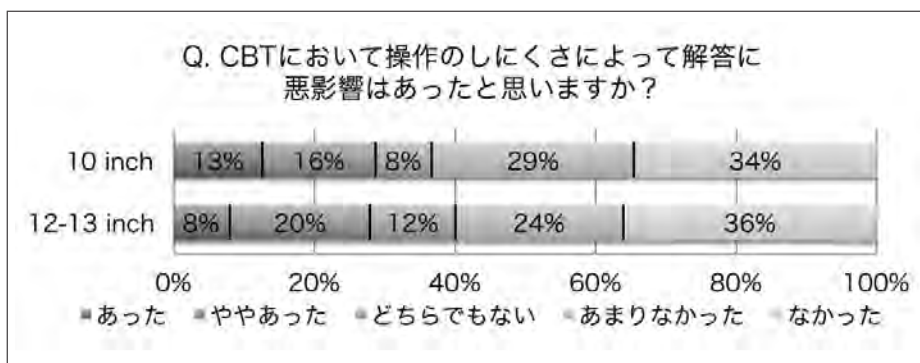


図 1.16: Q. CBT において操作のしにくさによって解答に悪影響はあったと思いますか？

また、以下の2つの項目は、紙筆とCBTの両方で受験した者のみが解答したアンケート項目である。対象者が利用していたタブレット端末は12-13型の大型のタブレット端末である。

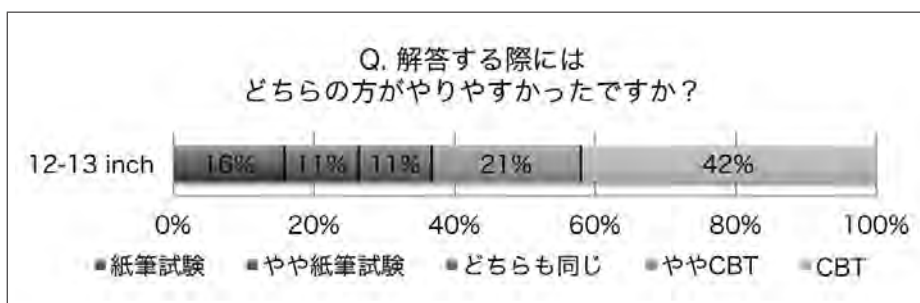


図 1.17: Q. 解答する際にはどちらの方がやりやすかったですか？

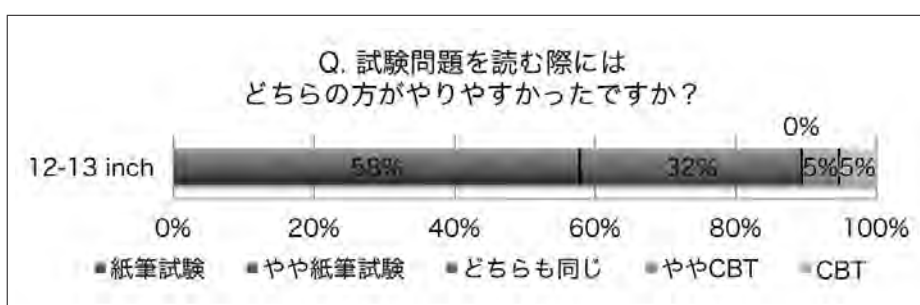


図 1.18: Q. 試験問題を読む際にはどちらの方がやりやすかったですか？

また、下記の表はアンケート結果の特徴的なものをまとめたものである。

表 1.15: Q. CBT では、1 問ごとに時間制限が設定されていました。1 問ごとに制限時間が設定されていることは、解答に影響がありましたか？

画面サイズ	選択肢	内容
10 inch	なかった	解答するのに十分な時間だったため。
10 inch	ややあった	合計問題数が分からないため時間配分が決められず、制限時間をその目安とすることもできないため。
10 inch	あった	時間のかかりそうな問題をとばすこと(あとまわし)ができない
12-13 inch	ややあった	国語など、長い資料がある場合には質より速度を意識させられる。完答が第1条件だから。
12-13 inch	ややあった	キーボード操作のミスにより、解答が消去されてしまった際、制限時間に到達してしまい再度解答することができなかった
12-13 inch	ややあった	時間制限があると、ひとつの問題にこだわりすぎず妥協するようになった。
12-13 inch	あった	時間が足りず中途半端な解答。
12-13 inch	あった	全体の試験時間との時間配分とそれぞれの問題に対する時間配分の兼ね合いが難しかった。
12-13 inch	あった	8割がた正解であろう解答をまず書いて、時間が余れば戻って100%正解の解答となるよう解き直す、ということが出来ないのはとても不便に思います。

表 1.16: Q. CBT において、キーボード入力によって回答に悪影響はあったと思いますか？

画面サイズ	選択肢	内容
10 inch	ややあった	タイピング能力も同時に見られ、そこを考慮しないとしたら不利な人が出てくる。
10 inch	ややあった	予測変換によって覚えていない用語も入力することができる。
12-13inch	なかった	むしろ訂正も容易にできてよかった。
12-13inch	あまりなかった	強いて言うのならば、分からない漢字も変換してくれてしまうこと。
12-13inch	あまりなかった	手でペンを握って書くよりも楽に入力できた。
12-13inch	ややあった	また、予測変換で出にくい漢字の入力が厄介。
12-13inch	あった	キーボード操作になれている人とそうでない人に差ができてしまう。

表 1.17: Q. CBT において、画面の見にくさによって回答に悪影響はあったと思いますか？

画面サイズ	選択肢	内容
10 inch	どちらでもない	拡大縮小できるので特に問題なし。
10 inch	ややあった	光が反射して資料が読みづらかった。
10 inch	あった	課題文に直接印をつけておくことができないため、文章の要点を書き出すか暗記する必要があるため。
12-13inch	なかった	画面は見やすかった。
12-13inch	あまりなかった	文章と図が離れている場合スクロールを要し、一度に両方を見渡せない。
12-13inch	ややあった	文章をよむとき少し横にもスクロールしないとよめないところがあった。
12-13inch	ややあった	資料と問題を全体的見比べられないのが不便。長文をスクロールで読むのは読みづらい。縮小すると文字が消えたりする。
12-13inch	あった	とても目が疲れる。人によってはバックライトを調整したい人もいるのでは。

表 1.18: Q. CBT において、操作のしにくさによって回答に悪影響はあったと思いますか？

画面サイズ	選択肢	内容
10 inch	なかった	パソコンの操作は慣れていたので。
10 inch	どちらでもない	慣れていない人には操作が難しいと思う。
10 inch	ややあった	キーボードが小さいので誤入力しやすい。
10 inch	ややあった	いつも自分が用いているのと少しキーボードの形態が違ったため
10 inch	あった	ページを変えづらい。
12-13inch	なかった	操作しやすかった。
12-13inch	なかった	練習問題で慣れることができたから。
12-13inch	あまりなかった	事前にインストラクションの紙と練習問題ができた点がよかった。
12-13inch	あった	スクロールしにくい。

表 1.19: 自由記述

画面サイズ	内容
10 inch	50 字以上になると全てを同時に見れない。
10 inch	センター試験のように、数十万人が受験する試験だと記述式の問題を 1 人の採点者が 1 つの問題を見ることが不可能なわけで、その点で不平等が生じるのではないかと感じた。解答の訂正が不可能で、見直しができないという点も改善すべき点だと感じた。
12-13inch	予測変換で先に単語が出てきてしまうことがあったため、重要な単語も漢字等はアバウトに覚えておけば入力できてしまいました。
12-13inch	漢字を忘れていても「変換」ボタンを押せば出てきてしまう。PC に慣れていないとかなり不利。全体の問題数や内容を見てから自分のやりやすい順番に解くことが不可能であった。
12-13inch	解答をする際はコンピューターのほうが修正もしやすく、また速く記入することができるため楽であった。その一方で問題を読む際はやはり紙のほうが下線を引いたり、印をつけたりしやすい上に資料をあちこち見ることができると思った。また、コンピューターはどうしても画面なので目が疲れてくる。さらに一問ごとに時間配分がなされていると時間をかけたい問題にかけられなかったり、わからなくてパスしたい問題の時間が余ってしまったりするので、全体で時間が決められているほうが良いと思った。
12-13inch	机に対し垂直に立てる方式では後ろから前の人の画面がよく見えるため、カンニングが起きやすいのではないかと。書けない漢字も入力できるため、今まで以上に漢字を書く力が落ちるのではないかと。コンピューターは複雑な機械である以上、不具合が発生する可能性は高い。不具合が発生すれば受験生は動揺するだろうと思われる。とはいえ、コンピューターは今の生活で必須であるから、コンピューターによる試験をやる価値はあると思う。全ての問題を解答後、紙なら見直し、解き直しができるが、今回はそれができず、何もすることができなかった。この点については改良の要ありと認む。

2 採点支援による効率化の調査

2.1 採点効率化検証の目的

本研究では、大規模試験における記述解答形式の採点効率化のためのいくつかの技術について評価をおこなうものとする。具体的には、以下の2点について記述解答答案に対する採点作業にその技術を組み入れ評価をおこなう。

1. 従来の光学的な読み取り情報のみに依存するICR(Interigent Character Recognition; インテリジェント文字認識)¹から、光学的な読み取り情報に言語モデルを組み入れることにより、より高精度に手書き文字の認識を可能にする技術。
2. 採点者がコンピュータ端末上で答案の採点ができるようにするためのインターフェイスを備えた採点システム。

さらに、これらに加えて不完全ながらも得られた答案テキストデータをもとに、類似度の高い答案を凝集させるクラスタリング技術も用いることで、より採点者に低負荷な採点環境を提供することを可能にする方式について評価をおこなう。

2.2 実験の概要

評価実験は下記の要領で実施された。

表 2.1: 評価実験の概要

項目	内容
実験日時	2016年3月10日 09:30-17:30 (12:00-13:00を休憩とした)
採点者特性	採点業務の経験がある大卒・短大卒の者
採点対象の問題	50字以上80字以内での記述解答を求める問題
実験群採点者数	13名
対照群採点者数	13名

採点実施風景は以下のようなものであった²。

¹活字認識のためのOCR(Optimal Character Recognition)ではなく、手書き文字認識の技術を指す

²採点実験時には、別の試験問題の採点作業を目的とした作業従事者もいた。



図 2.1: 採点実施風景

また、実験群と対照群において、それぞれ次のような体制のもとで採点作業をおこなった。

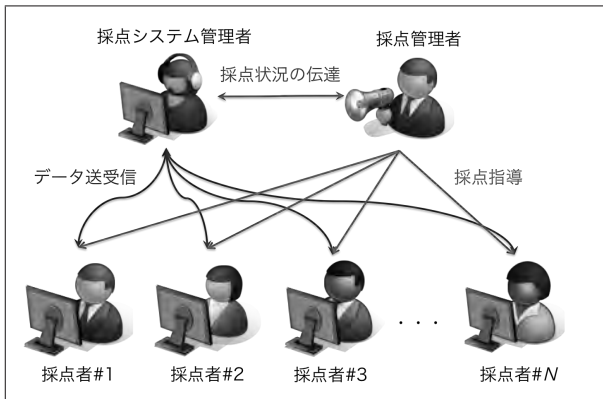


図 2.2: 実験群の採点管理体制

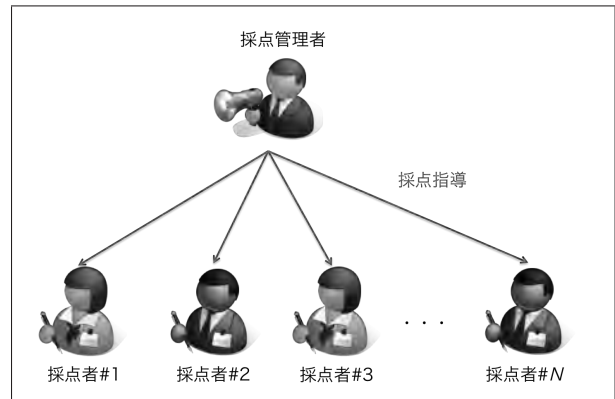


図 2.3: 対照群の採点管理体制

2.3 実験の手続き

実験群と対照群それぞれにおける実験手続きは以下のとおりであった。

1. 実験群

- (a) 50 万人分の手書き答案に対して、言語モデルを組み入れた ICR をかけることで、テキストデータ化をおこなう。
- (b) 不完全ながらもテキスト化された答案データに対して、クラスタリングをおこない、1 万クラスに答案を分類する。

- (c) 1万クラスタの中から、ランダムに6クラスタを選び、その中から500人分の答案を抽出する。
- (d) 縦書きである500人分の答案をマス目ごとに画像切り出しをおこない、横読み用に並び替える。
- (e) コンピュータ端末にクラスタリングされた500人分の横読み答案画像を順に呈示しながら、採点をおこなう。
- (f) 採点は観点ごとにおこなっていく

2. 対照群

- (a) 約50万人分の手書き答案の中から、ランダムに500人分の答案を抽出する。
- (b) 縦書きである500人分の答案をマス目ごとに画像切り出しをおこない、横読み用に並び替える。
- (c) 答案画像をコンピュータ端末上での表示用と等倍で紙に印刷をする。
- (d) 紙に印刷された答案画像を見ながら、記録用紙に採点結果を記録していく。
- (e) 採点は観点ごとにおこなっていく。

2.4 実験に用いた試験問題

なお、用いられた問題について下記に示す。

【フリップ】

1 オカリナとは

- ハトに似た形の笛
- イタリア生まれ

イタリア語では「小さなガチョウ」

2 オカリナの特徴

吹き口 音色は？

材質は？ 指穴の数は？

3 気軽に楽しめる

- リコーダーと似た指使い
- 持ち運びやすい大きさ
- 簡単な曲ならすぐ吹ける！！

【ノート】

発表の流れ	その他の情報
<ul style="list-style-type: none"> ■ 閉き手を見てから話し始める <input type="checkbox"/> 最初の言葉 	<ul style="list-style-type: none"> ※ 状況に応じて使用する
<p>1</p> <ul style="list-style-type: none"> <input type="checkbox"/> オカリナはハトの形に似た楽器。 <input type="checkbox"/> イタリアで生まれ世界に広まった。 <input type="checkbox"/> イタリア語では、「小さなガチョウ」という意味である。 	<ul style="list-style-type: none"> ※ ガチョウの頭の形に似ている。体の様々な観がある。
<ul style="list-style-type: none"> ■ 開いけて間を置く 	
<p>2</p> <ul style="list-style-type: none"> <input type="checkbox"/> 材質はフー糸織きの陶器のものが多い。 <input type="checkbox"/> 音色はフー吹き口から息を吹き込むと、温かみのある音色が出る。 <input type="checkbox"/> 指穴の数はフー12個が多い。(表側に10個、裏側に2個) 	<ul style="list-style-type: none"> ※ プラスチックのものもある。 ※ イタリアでは、10個が多い。
<p>3</p> <ul style="list-style-type: none"> <input type="checkbox"/> リコーダーと同じような指使い。 <input type="checkbox"/> ポケットに入るくらい大きさ。 <input type="checkbox"/> 簡単な曲なら、すぐ吹けるようになり、吹ける曲が増えると、ますます楽しくなる。 	<ul style="list-style-type: none"> ※ いろいろな形や大きさのオカリナがある。
<ul style="list-style-type: none"> ■ 質問がないか尋ねる <input type="checkbox"/> 終わりの言葉 	

森田さんは、国語の時間に、自分の特技についてスピーチをする学習に取り組んでいました。次は、発表の際に使用する「フリップ」と「ノート」です。これらを読んで、あとの問いに答えなさい。「フリップ」と「ノート」にある①から③は、発表の際に使う順番を先します。

①「フリップ」の表側に用いる。②～④は裏側に用いる。⑤～⑦は裏側に用いる。

図 2.4: 実験に用いた試験問題 (平成 27 年度全国学力テスト国語 B 問題より引用)

2.5 採点観点と採点基準

上記の実験用問題において、採点においては下記の二つの採点観点と、それぞれの観点における採点基準について下記に示す。

2.5.1 観点1

【ノート】の内容と結び付けて、〈演奏するタイミング〉を選んだ理由を具体的に書いている。

1. 観点1において、正解とするものの例

- (a) 「温かみのある音色といっても言葉だけでは分かりづらいと思うので、実際に演奏して聞かせれば、聞き手が音の温かみを実感し、一層興味をもってもらえると思うからです。」(79字)
- (b) 「フリップ2のときにオカリナを吹くところを見れば、吹き口への息の吹き込み方や、指穴を押さえる様子がよく分かり、興味をもってもらえると思う。」(69字)
- (c) 「私なら3のときに吹きます。なぜなら、ここで吹けば、オカリナがリコーダーと同じような指使いであることを見せられ、演奏の仕方を理解してもらえるからです。」(75字)
- (d) 「実際に簡単な曲を吹いてみせれば、この曲ならすぐに吹けそうだと思います、聞き手がやってみたい気持ちになるから。」(53字)

* なお、観点1について、〈演奏するタイミング〉で使用する【フリップ】の直前や直後の【ノート】の内容と結び付けて書いているものも正解とする

- (e) 「フリップ1でオカリナの形について説明しているので、フリップ2の始めに実物を見せながら演奏すると、本当にハトの形に似ていることが分かるので興味がもてる。」(76字)
- (f) 「2の終わりで吹きます。次にリコーダーと同じような指使いであることを説明するので、吹くところを見せておけば、自分もやってみたいと興味をもってもらえると思います。」(80字)

* なお、観点1について、〈演奏するタイミング〉で使用する【フリップ】以外の【ノート】の内容も含んで書いているもの(例示の下線部は関係箇所)も正解とする

- (g) 「3で演奏すると、リコーダーと同じような指使いであることが分かるし、指穴の数も見せながら分かりやすく説明できるから。」(58字)
- (h) 「オカリナについて説明するときに、実際にオカリナを見せて演奏する方が聞いていて楽しいと思うからです。」(50字)【ノート】の内容と結び付けずに〈演奏するタイミング〉を選んだ理由を書いている。
- (i) 「2の材質や音色の説明のときに吹くよりも、吹ける曲が増えると楽しくなるという説明のときにオカリナを吹いた方がよいから。」(59字)〈演奏するタイミング〉を選んだ理由を書いていない。
- (j) 「聞き手に実物を見せることで、オカリナがハトやガチョウの形に似ているということを確かめてもらえるので、私はフリップ3の説明のときに演奏します。」(71字)〈演奏するタイミング〉と選んだ理由がずれている。

- (k) 「3のときに演奏することで、リコーダーと同じような大きさだということがよく分かり、ポケットに入るくらいの大きさだと思ってもらえるから。」(67字)【ノート】の内容を正しく理解していない。

2. 観点1において、不正解とするものの例

- (a) 「オカリナについて説明するときに、実際にオカリナを見せて演奏する方が聞いていて楽しいと思うからです。」(50字)【ノート】の内容と結び付けずに〈演奏するタイミング〉を選んだ理由を書いている。
- (b) 「2の材質や音色の説明のときに吹くよりも、吹ける曲が増えると楽しくなるという説明のときにオカリナを吹いた方がよいから。」(59字)〈演奏するタイミング〉を選んだ理由を書いている。
- (c) 「聞き手に実物を見せることで、オカリナがハトやガチョウの形に似ているということを確かめてもらえるので、私はフリップ3の説明のときに演奏します。」(71字)〈演奏するタイミング〉と選んだ理由がずれている。
- (d) 「3のときに演奏することで、リコーダーと同じような大きさだということがよく分かり、ポケットに入るくらいの大きさだと思ってもらえるから。」(67字)【ノート】の内容を正しく理解していない。

2.5.2 観点2

五十字以上、八十字以内で書いている。

1. 字数を満たしていない。

2.6 結果

計測された時間から、1答案あたりの採点時間を算出したものを下記の図に示す。それぞれ、観点1と観点2について、実験群と対照群が採点した時間の平均を示している。

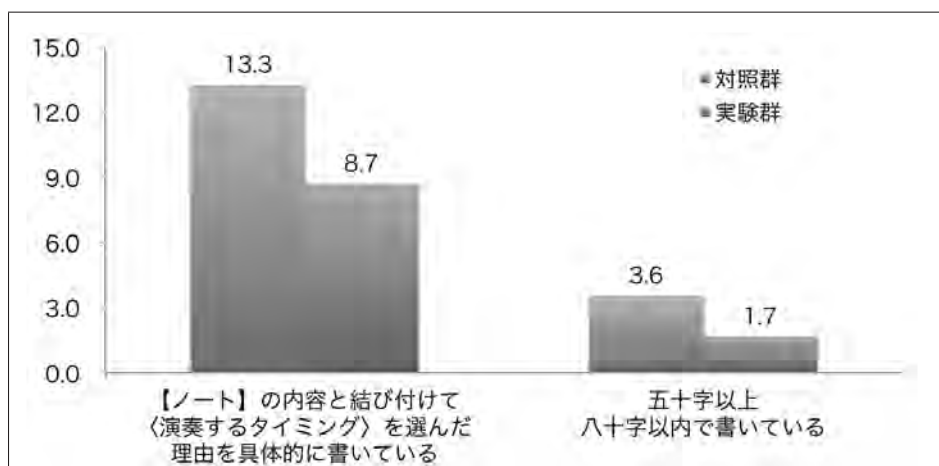


図 2.5: 実験群と対照群の 1 答案あたりにかかる採点時間 (秒)

結果から、クラスタリングをかけ、さらにインターフェイス上で採点したグループは観点 1 において約 8.7 秒、観点 2 において約 1.7 秒であった。一方、ランダムに答案を手書きで採点した群は、観点 1 において 13.3 秒、観点 2 において 3.6 秒であった。これより、ランダムに採点用紙に記録していく群に比べて観点 1 においては 35% 程度短縮され、観点 2 においては 50% 程度短縮されるという結果が得られた。

2.7 考察

本実験において、言語モデルを組み入れた ICR に基づくクラスタリングを援用したり、採点インターフェイスを用意することで採点時間の短縮が見られることが明らかになった。

採点時間の短縮の割合については、下記の要素によって変わる。どんな記述解答問題や採点基準でも今回のように 30% 超の時間短縮が得られるわけではないことに注意が必要である。

1. 制限文字数：制限文字数が多くなると、文と文のつながりや文章の構成が複雑になることになり、答案の多様性が広がる。そのため、文章全体の中の一部だけが類似している文章同士を並べても、文章間での相違のある箇所が増えいき、結果、類似性の高くない可能性が高まると考えられるため。
2. 試験問題の内容 (文章中の語を用いさせるか否か)：試験問題中の文章から一部を抜き出させるような問いであると、答案に用いられる語の種類が限られてくるため、より自然言語処理技術による答案の弁別が正誤判定につながりやすくなると考えられるため。
3. 採点基準：機械的な判断が可能なものは、クラスタリング等の技術を用いることで短縮につながると考えられる。一方、「自分の経験を例にあげて、あなたの考えを書きなさい」というような、出てくる語が文章の特徴判断に直接的に結びつかないような文章を書かせると、採点短縮にはつながりにくいと考えられる。

したがって、現段階において「採点を効率化する」という点において効果のあるものは

1. 制限文字数を短く設定する。
2. 試験問題文中に出てくる語を用いさせるなど、文章に現れる語の種類を少なくする。
3. 採点基準は、解答の表層的な部分について採点者が機械的に判断していけるようなものにする。具体的には、キーワードが入っているか否か、文字数制限を満たしているか否か、特定の言い回しを用いているか否か。

などであろう。これらを満たすような答案であれば、思考力・判断力・表現力を構成する諸能力を中心に評価することに留意しつつも、採点支援のための技術は効果を発揮していくであろう。

一方、多くの文字数を書かせるような論述解答形式や、採点基準が複合的で複雑なものなどを効率化させるためには、更なる技術や理論の発展が必要になるであろう。

2.8 実験群の採点実施風景

ここでは、実験中の採点実施風景等について示す。



図 2.6: 実験群の採点風景

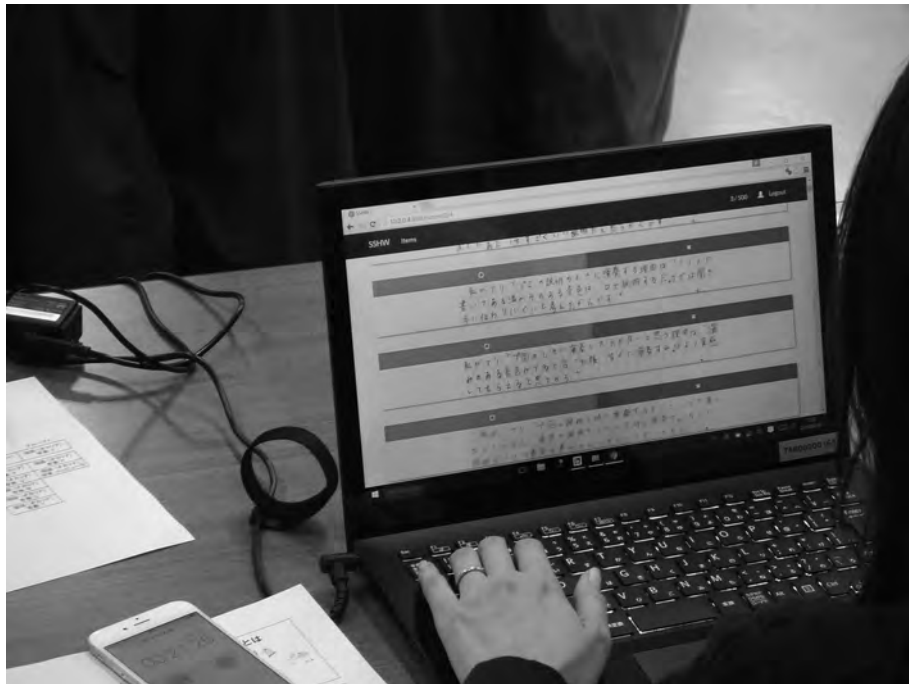


図 2.7: 採点インターフェイス

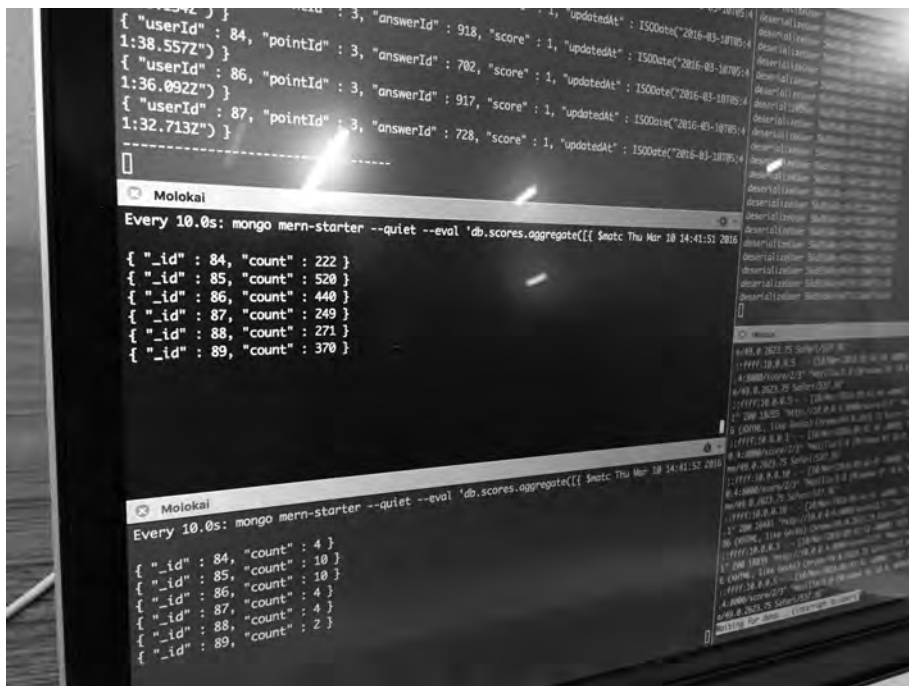


図 2.8: 採点システム管理者画面

2.9 付録

2.9.1 実験刺激に用いられた答案画像とデータ

ここでは、採点の対象となった答案画像をテキスト化したものを実験群と対照群に別けて示す。さらに、実際に採点で用いた答案画像データについても示す。

表 2.2: ランダムに抽出された答案テキストデータ (一部)

テキスト化データ
まず、オカリナの説明をして、その次に、声力ソナの特徴の説明をして、最後に、気軽に楽しめることを説明して、簡単な曲ならすぐ吹けると言った後に演奏した方が” から吹き口から息を吹き込むと温かみのある音色が出る。とあるのでそこで音色を聞かせるといいと思う。次に簡単に吹けると言うので、その前にやつた方がいいと思う。フリップ回ではオカリナの音色について説明しています。それでーートたは温かみのある音色が出るを書いてあるので、この時に演奏すればよりその音を実感できただけです。演奏するタイミングは、図の説明のときが良いと思う。理由は、図では、オカリナについて説明しているし、問いの中にも音色について聞いており、答えが分かりやすいから。ハ主からオカリナを吹きたいと思いますのでの前にオカリナの説明をします。オカリナの材質は素焼きの陶器のものが多くひす音色は吹き口から息を吹き込むと温かみのある音色が出る20所吹き。から息を吹き”ぞと温かみのある音色が出る”と説明して”スピーチの聞き手がその音色に興味を持”た所で演奏した方がいいと思、だから、”フリル5回の説明の7温かみのある音色が出るレというのとフナリップ回のつりコーダーのような指づかいレという二つの説明を演奏で同時に説明できるからであるヨの説明のときに、オカリナを実際に演奏したほうがいいと思います”なぜなら、音色を口でご説明するよりも、実際に演奏したほうが、音色かわかりやすいからです。回の説明の時に演奏をす3かと言うとフリップのイラクのの横に、イ音色はマレと書かれていたので、その時に吹けばくちし説明でフヲと思ったからです。フリップ回の説明の時に、音色の説明があるので、そこで一曲演奏す3と聞き手は温かみのある音色とはどんなものかわかるので、興味をもってもらえると思うから。フリップ図でオカリナの材質や音色、指穴の数について説明しているのでフリップ回の説明のときに演奏するとオカリナについてわかりやすいと僕は感、たかぢです。フリップ回の説明のときなら、聞き手が疑問をもちそうな材質や音色、指穴の数などの特徴をオカリナを演奏し井回の説明をすればイメージがしやすいと思うからです。ノードの回では、オカリナの特徴について説明していく回では、オカリナの吹き方などを説明していく一つフリップの回でもオカリナの吹き方などがあるから。回の説明のときにつ温かみのある音色が出るレと説明したものの、実際にどのような音色が出るのか聞き手は知らないため興味をわき、聞いてみたいと思うから回のオカリナの説明が終わってからのほうがオカリナについてよく分かっているし、図の所にある気軽に楽公ぬる所を見てもらった任うが分かりやすくなるから回のYきレオカ”ナへ特徴でご村井刃升お音色・材質、指穴の数など具体的に説明しべ”るので僕だ、たらオカリナの特徴をした後の図のタイミこ々がいい乞思う吹き口から息を吹き込むと温かみのある音色が出る、と”う説明をしたときに実際に聞き手側に吹いても5じし指穴の数を説明した時に指穴をおさえながら吹いてモ5うかご。フリップ3のときに演奏することによって、つこのようにレとフリップ3に書いてある内容をあとから付け加え”更に演奏と結び付けた発表をすることが出来るから。なぜ、フリップ2のときし演奏する理由はノートの所で、温かみのある音色が出てくるが、実際温かみのある音色がどんなものか聞いてまらうため。これを選んだ理由ば回の説明の時ほとんどな音色かを説明(ます”なのでこの時に実際にたど人な青がす3のかを聞いても”たほちがわかりやすいと考えたかぢです。僕は、Aのタイミングで演奏した方がいいと思います”なぜなら、回の時に、音色の説明をしていても時に聞き手に音を伝えた方が印象に音が残りやすいし見なたからです。理由は、二のフリップは、オカリナの特徴だけを説明して、三のフリップは、オリリナの使い方などを説明して、フ簡単な曲ならすぐ吹けるハし後に演奏したらいい。ソリップ回の説明のときに、音色を説明す3とニ3があるので”その時に演奏すれば、自分の目で見たり感じたりすることができるのでクリップ回の説明のときたしました。へ珀き。から息を吹き込むし、温かみのある音色が出るいし言”た彼に、つ申際に吹”てみキシ”三だじし言って吹”て83し興味を持”てキぢえると思うたかぢ。オカリナを演奏するタイミングは回の説明のときが良”と思う”理由は”回では材質や音色について説明し”イ、ここで演奏すれば、材質や音色についてよく分かるから。相手からつ音色はアレと聞いかけられているのに対し、フ吹き口から息を吹き込むと、温かみのある音色が出るしと解答しているため実際に吹くのに良いと思ったからです。図の説明のときし音色のことを説明しているが、図の説明トは指仗”や曲のことを説明しているから。たから演奏をするならば、田の説明のときの方が戻”と思う。回の説明のときに演奏する理由は回の説明で演奏のことについて話しているからです。話していきと演奏すれば一聞き手も興味をひかれると思いきした。フリップ2レは、オカリナの音色について説明しているが、言葉だけでは想ぞうしにくいと思ったから。また、フリップ3の説明にもつながるためより、理解しやすいから。フリップ回のときに、音色は吹き口から息を吹き込むと”温かみのある音色が出る”と説明しているの、その時、実際に音色を聴かせてあげたたちがイメージしやすいと思うから。

表 2.3: クラスタリングされたテキストデータ (一部)

テキストデータ
オカリナは、リコーダーと同じような指使いで簡単な曲ならすぐ吹けるようになるからフリップ回の説明をしている時にオカリナがどの上に乗しめるのか聞き手が見たいから。
オカリナは、リコーダーと同じような指使いで簡単な曲ならすぐ吹けるからオカリナを実際に演奏したら聞き手のハト興味を右つてもらえるから。
オカリナは、リコーダーと同じような指使いで簡単な曲ならすぐ吹けるようになるというのを証明するトよじつフリップ一回の流叩かきき吹いき大が良から。
オカリナは、リコーダーと同じような指使いで簡単な曲ならすぐ吹けるから江ろから4いう説明のしきに演奏する方が指使いなども見れて聞き手に共感を得やすいから。
オカリナは、リコーダーと同じような指使いで簡単な曲ならすぐ吹けると告げてあるので、実際にオカリナの演奏を聞けば、モフとオカリナに興味を持つと思うから。
オカリナは、リコーダーと同じような指使いで簡単な曲ならすぐ吹けることを説明したうえで実際に演奏すれば、聞き手は説明した点に注意してより興味をもつと思うから。
オカリナは、リコーダーと同じような指使いで簡単な曲ならすぐ吹けるという事を立証すため因の説明のときが一番いいタイミングだと思います。
オカリナは、リコーダーと同じような指使いで簡単な曲ならすぐ吹けることもできて、実際にみんなの前で吹けばよりオカリナについてみんなが興味を持つと思ったから。
オカリナは、リコーダーと同じような指使いで、曲かすぐ吹けるようになるというツイかいはほうがただ吹イろたけなようにかえてしまうかべ。
オカリナは、リコーダーと同じような指使い、簡単な曲なら、すぐ吹けるようになるなどと説明をしてから、実際にオカリナを吹けば、さらに分かりやすくなるかぶ。
オカリナは、リコーダーと同じような指使いだということと簡単な曲ならすぐに吹けるように打ると説明するので、ここで吹けば見ている人に伝わりやすいと思うから。
オカリナは、リコーダーと同じような指使いで、簡単な曲なら、すぐ吹けるからユのタイミングでオカリナの演奏をしてあげれば、リコーダーのような指使いなども分かるから。
オカリナは、リコーダーと同じような指使いで、簡単な曲なら、すぐ吹けるようになることを説明した後の方が、聞いている人が理解しやすくなると思うから。
オカリナは、リコーダーと同じような指使いで、簡単な曲なら、すぐに吹けるようになってすぐ楽しめることを教えてからの方が、自分でも出来るかしれだと思えるから。
オカリナは、リコーダーと同じような指使いで、簡単な曲なら、すぐに吹けるようになるということがわかった。それに持ち運びやすい火きさということもわかった。
オカリナは、リコーダーと同じような指使いで、簡単な曲ならすぐ吹けると説明した後いんば、叶下なも興味をもつてきり2くしると思うから。
オカリナは、リコーダーと同じような指使いで、簡単な曲ならすぐ吹けるようになるという事を説明した後に実際に吹けるといふ所を見せた方が良かったと思うから。
オカリナは、リコーダーと同じような指使いで簡単な曲ならすぐに吹けるようになることと先に言うことで実際に演奏を聴いてみて、より興味を持ってくれると思つたから。
オカリナは、リコーダーと同じような指使いで簡単に吹けると説明しているの、このタイミングで演奏すればオカリナが簡単さがより伝わると思うからC
オカリナは、リコーダーと同じような指使いで1簡単な曲なら、すぐ吹けるようになるという説明から、本当にすぐに吹けるのかを見せた方がいいと思うから。
オカリナは、リコーダーと同じような指使いで簡単な曲なら、すぐ吹けるようになるという事を説明してからのの方が聞き手が自分も吹けるかもと思うと思つたから。
オカリナは、リコーダーと同じような指使いのため、簡単な曲ならすぐに吹くことができるという説明のあるフソツろの時に吹けばみんなに興味を持ってもらえると思うから。
オカリナは、リコーダーと同じような指使いで、簡単な曲なら、すぐ吹けるようになるという事を説明した後に演奏したら、聞き手に興味を持ってもらえると思つたから。
オカリナは、リコーダーと同じような指使いで簡単な曲ならすぐ吹けるようになるという事を実際に演奏して理解させる。そうすれば興味わいてくるY思つたカウである。
オカリナは、リコーダーと同じような指使いで簡単な曲なら、すぐに吹けるようになるという事を説明して演奏すれば、聞き手もオカリナに興味を持つと思うから。
オカリナは、リコーダーと同じような指使いで、簡単な曲なら、すぐ吹ける、と言って、ねの乙簡単な曲をそこで一曲吹けば、聞き手に一層興味をもってもらえるから。
オカリナは、リコーダーと同じような指使いで、簡単な曲ならすぐに吹けるようになることを、自分が見せると、みんながやってみたくて興味をもてくれると思。たから。
オカリナは、川日ハ日と同じような指使いで簡単な曲なら、すぐ吹けるようになると言て、自分か吹くこYができる曲を発表した方ケから。
オカリナは、持ち運びが簡単にひきて誰でも簡単だ曲ならすぐ吹けるようになるから、また、練習し吹けス曲が増すと楽しくだ2から。
オカリナは、リコーダーと同じような指使いで、簡単な曲はすぐ吹けるようになるという事を説明してからのの方が聞き手が自分も吹けるかもと思うと思つたから。

2.9.2 実験従事者

本実験を遂行するに際しての実験従事者は以下のとおりであった。

1. 京都大学・黒橋禎夫教授：言語モデルを組み入れたICRの実施，クラスタリングの実施
2. 東北大学・乾健太郎教授：ICR，クラスタリングの技術検討
3. 大学入試室(富士通株式会社)：約50万人分の答案画像の取得，答案画像からのテキスト化，縦書き答案画像のマス目ごとの切り出しと横読みへの変換
4. 大学入試センター・大久保：採点インターフェイスの開発と採点実験の実施

c-2 タブレットコンピュータを用いた

障がいのある者のための受験上の配慮方法の検討

1 概要と背景

平成 32 年度に予定される大学入学者選抜の改革, 特に「大学入学希望者学力評価テスト(仮称)」における障害のある受験者への配慮を実現するための基礎的研究として, タブレットコンピュータを用いた任意倍率による画面表示と表示内容の音声読み上げを付与した試験問題提示システムの開発を推進した。

障害を有する受験者, 特には弱視者や読字障害者は, 問題冊子上の文字を視認する方法での閲覧に大きな制約を有する。「大学入学希望者学力評価テスト(仮称)」導入に際しては, これを改善する受験上の配慮が法的・社会的に要求される(障害者差別解消法に規定された「合理的配慮」義務)。そのための出題システムとして, タブレットコンピュータを用い任意倍率の拡大表示や音声読み上げを実現したものが有望である。本研究では, 今後の開発のための論点の抽出・整理を目的として, センター試験過去問の表示と読み上げを実現するシステムの試作を行った。なお, 本研究が扱うタブレットコンピュータを用いた提示システムは, 先端的な CBT を実現する出題システムとは性質を異にするところも多い。むしろ, いわゆる電子書籍閲覧アプリに類するソフトウェアを基礎に, 障害者の利用の利便性を追求するとともに, 試験問題閲覧に特化させたものという理解が妥当である。

上記を踏まえ, 研究期間中, システム(タブレット用試験問題閲覧・読み上げアプリ)の要件決定, 開発, 評価を行った。当初示されたニーズのうち, 開発プラットフォーム上で実装可能なものは, ほぼ実現した。

- 拡大表示については, 弱視者教育で指標とされている各種拡大倍率(18 ポイント, 22 ポイント, 26 ポイント)に網羅的に対応できた。
- 加えて, 画面表示に同期し, 読み誤りのない試験問題の音声読み上げを実現した。

開発したシステムは, 閲覧のユーザビリティや運用上の安定性を評価するためのモニター調査に供する段階に到達している。

他方, 漢文など一部の試験問題はそのままの形態では出題できないことや, 音声読み上げ速度改善の必要, タブレット管理・運用体制検討の必要が判明した。研究という観点から, 適宜対応策を定式化した。本研究では, もっぱら試験問題提示手段の確立を課題とする。視覚障害や発達障害のある受験者は, 読み書きにしばしば困難を抱えるため, 試験時間延長の必要性やその延長倍率の産出が重要であるが, 本研究では取り扱わない。こうした課題は, 点字冊子出題や拡大文字冊子出題など, 他の受験配慮との間で横断的に存在する課題として今後の検討が必要である。

1.1 「大学入学希望者学力評価テスト(仮称)」における障害のある受験者への配慮の位置づけ

中央教育審議会答申「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」(平成 26 年 12 月 22 日)において, 現在の大学入試センター試験に代わり「大学入学希望者学力評価テスト(仮称)」(以下「学力評価テスト」)を導入することが提言された。同答申では, 新しい時代の入試として, 多様な背景を持った学生を受け入れられるような選抜方法の重要性が繰り返し強調されている。障害のある受験者への配慮に関しては, 「入学希望者の経済的負担や受検場所, 障害者の受検方法を考慮するなど, 受検しやすい環境を整備する」と述べられている。つまり, 障害受験者がこ

れまでより受験をしやすくなるような配慮の充実が必要な一方で、受験者に経済的負担を負わせないことが求められる。

1.2 障害のある受験者への配慮が必要とされる法的・社会的背景

1.2.1 「合理的配慮」の義務化

国連障害者の権利条約を具体化する国内法である障害者差別解消法が、2016年4月より施行される。当該条約・法律の特色は、「合理的配慮」の概念を導入し、障害者の高度な社会参加の実現を目指すところにある。合理的配慮とは「障害者が他の者と平等にすべての人権及び基本的自由を享有し、又は行使することを確保するための必要かつ適当な変更及び調整であって、特定の場合において必要とされるものであり、かつ、均衡を失した又は過度の負担を課さないもの」（国連障害者の権利条約、第二条）と定義され、公的機関にはその履行が義務付けられる。

1.2.2 大学入試に関わる障害受験者配慮充実の動向

前述の法的基盤を具体化するために、文部科学省は2015年に「文部科学省所管事業分野における障害を理由とする差別の解消の推進に関する対応指針」を策定した。この指針の中では「ルール・慣行の柔軟な変更の具体例」として、「入学試験において、本人・保護者の希望、障害の状況等を踏まえ、別室での受験、試験時間の延長、点字や拡大文字、音声読み上げ機能の使用等を許可すること。」が配慮事項として明示されている。このことは、公的な入学者選抜を実施する組織や国公立大学には、列挙された配慮を行うことが義務づけられることを意味する（私立大学は努力義務）。

他方、大学入試センター試験においては、近年いくつかの新しい受験上の配慮が行われている。

1. 平成27年度センター試験（2015年1月実施）では、発達障害により読みに困難を抱える受験者に対して、試験監督者による問題文の代読を認めた。これは個別対応事例であるが、文字メディアでの閲覧を前提に作題された試験問題を音声メディアで提示したという点で重要である。
2. 平成28年度センター試験（2016年1月実施）より弱視者や一部の発達障害者への配慮として、22ポイントの拡大文字を用いた問題冊子による出題を実施した。この問題冊子には用いられる文字サイズが元冊子の2倍を超える。そのため、単純拡大コピーではなく、再編集を伴う製版・印刷となる。問題冊子の総ページ数は元冊子のそれから増加する。

1.3 出題方法に関わる配慮の具体的な実現手段

「大学入学希望者学力評価テスト（仮称）」においては、上記の背景を踏まえた障害のある受験者への配慮実施が要請される。具体的には、「文部科学省所管事業分野における障害を理由とする差別の解消の推進に関する対応指針」に示された点字や拡大文字、音声読み上げ機能を使用し、安定的に実施・運用できる受験手段の確立が喫緊の課題である。

上記3種類の受験手段のうち、点字による受験に関しては、点字冊子問題を用いた出題と点字解答用紙による文字解答が共通第1次学力試験導入時より用いられてきた。重度視覚障害児初等中等教育の動向を勘案しても、「大学入学希望者学力評価テスト（仮称）」では点字冊子を用いた試験実施の踏襲を軸に検討する必要がある。

拡大文字を用いた冊子による出題に関しては、平成28年度センター試験により導入された

22ポイント冊子による出題が一つの基準事例となる。ただし、弱視者の視覚認知の特性は多様であり、弱視児初等中等教育においても、多様な拡大倍率と配色の変更を行った教科書や教材が用いられている。一律の拡大文字冊子による出題で十分な配慮が実現できるとは考えにくい。『特別支援学校及び通常の学校に在籍する視覚障害のある児童生徒の教科指導の質の向上に関する研究』（独立行政法人国立特別支援教育総合研究所，2010年）では、弱視時の初等教育における拡大文字利用について、以下の2点が報告されている。

1. 弱視児童生徒の利用に供するために作成される拡大教科書では22ポイントの拡大文字が標準的に用いられる。加えて、18ポイント版および26ポイント版がしばしば製作される。
2. 1.の文字サイズでは不十分な児童生徒のために、平均33ポイントから36ポイントの文字を用いた拡大写本がボランティアにより製作され利用されている。

他方で、印刷や校正の人的・経済的コストを考慮すると多様な拡大文字冊子を用いた試験実施を安定的に運用することには多くの困難が予想される。

「音声読み上げ機能」とは具体的にはタブレットコンピュータやパソコンを用いてコンピュータ上で電子データを音声読み上げさせる手段を示す。現状、国内ではこのような手段を用いて厳密な公平性を追求した試験の安定的運用は前例がない。平成27年度センター試験で実施された人による代読は「音声読み上げ機能」には該当しない。また、代読を「音声読み上げ機能」に代わる次善の手段として用いることは、受験者の問題閲覧・解答効率への大きな制約に加え、安定的な実施・運用が極めて困難であり選択肢とできない。他方で、支援技術の活用注目する近年の特別支援教育の動向を踏まえれば、「音声読み上げ機能」を用いた出題へのニーズは今後増加することが予想される。

参考—海外におけるコンピュータを用いた受験配慮

コンピュータを用いて試験問題を提示する受験配慮を全国規模の統一学力テストに導入した事例として、スコットランドのSQA(Scottish Qualifications Authority)が実施するDigital Question Paperが存在する。Digital Question Paperは2008年より運用されている。試験問題は文字データが埋め込まれたPDFファイルで提示され、受験者は任意倍率での単純拡大による表示とコンピュータ音声合成を用いた音声読み上げで試験問題を閲覧することができる。

Digital Question Paperについては、利用する受験者数の推移や運用コストなど多角的な観点からの検証が行われている。スコットランドを含む欧米では、読字障害者に対して人の朗読による教育支援がしばしば利用されてきた。この検証の中では、こうした朗読に頼る受験に比して、Digital Question Paperが自律的、効率的に試験問題に取り組むことのできる手段である側面が強調されている。

1.4 本研究の課題

以上から、「大学入学希望者学力評価テスト(仮称)」における障害のある受験者への配慮に関わる喫緊の課題として、任意の倍率の拡大文字を用いた問題の提示と問題内容の(コンピュータ上での)音声読み上げを実現することが浮上する。その際、大学入学者選抜に要求される公平性の担保も達成されなくてはならない(ここにいう公平性の担保とは、受験者の不利益とならない受験環境の保証、試験実施トラブルの回避、不正の防止を含む)。

1.4.1 タブレットコンピュータを試験問題提示システムに用いることの有効性

これらの課題を解決しうるものとして、タブレットコンピュータを用いた試験問題提示システムに注目した。以下、タブレットコンピュータの利用がこれらの課題解決に際して有する有効性を述べる。

1.4.1.1 拡大文字を用いた問題の提示

タブレットコンピュータ用 OS には標準でピンチアウトなどと称される画面拡大機能が実装される。加えて、しばしば障害者の利用に配慮し、「ユーザ補助」や「アクセシビリティ」と呼ばれる種々の機能が実装される。これらの機能の中には、より高機能な画面拡大や配色の変更機能が含まれる。

ピンチアウト機能やアクセシビリティ機能としての画面拡大は、いずれも OS 全般で作用する機能である。これらの拡大は、コピー機を用いた拡大複写と同種の単純拡大となる。そのため、拡大後の表示では、画面を左右上下にスクロールすることで初めて拡大前の画面表示全体を確認することができる。

これに対して、リフロー表示を用いた拡大では、拡大後の表示で横方向にあふれる文字を次行に送り表示する。ユーザは左右のスクロールをせずに、文章を閲覧することができる（行数が増加するため、縦方向のスクロール回数は増える）。試験問題コンテンツのデータフォーマットを適切に選択し、提示ソフトウェア（アプリ）側で個別の対応を行えば、リフロー表示など、文書構造を踏まえたよりきめ細やかな拡大表示も可能となる。

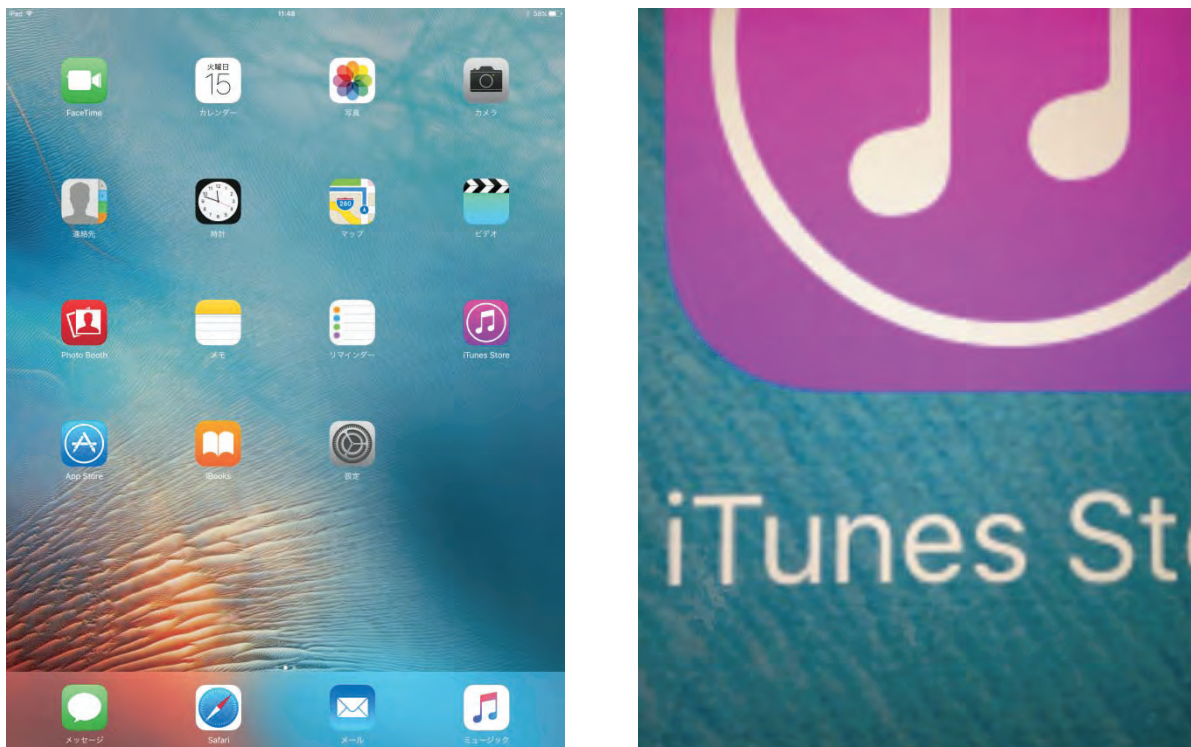


図1 アップルコンピュータ社 iPad Pro のホーム画面の標準状態での表示(左)とアクセシビリティ機能による画面拡大(右、拡大倍率最大 1500%)の表示

1.4.1.2 問題内容の音声読み上げ

現在流通するコンピュータには録音済みの音声データを再生する機能が備わっている(サ

ウンド出力)。また、多くのタブレットコンピュータ用 OS には、与えられた文字列を解析して音声データを生成する音声合成機能が実装されている。

1.4.1.3 上記2機能の同期

画面の表示内容を音声で読み上げる場合、表示と音声の連動が強く期待される。すなわち、現在読み上げられている箇所を画面上でハイライト表示して明示するような機能（ハイライト表示される箇所は、音声読み上げの進捗に応じて移動して行く）が望まれる。この同期はタブレットコンピュータのようなプラットフォームを用いる場合に実現できる重大なメリットであり、各種の支援技術においても広く実現されている。

1.4.1.4 大学入学者選抜に要求される公平性の担保

1.4.1.4.1 閲覧環境としての安定性

タブレットコンピュータは、機種選定や使用する機能の取捨選択を慎重に行えば高度な安定性を確保できるプラットフォームである。また、電子データの格納フォーマットに広く利用される標準規格を用いるなどの配慮を行えば、画面表示（試験問題の提示）の再現性は高度に保証される。こうした性質は、公平性を害するようなコンピュータトラブルや不具合を回避する上で大きなメリットとなる。

1.4.1.4.2 想定受験者の操作習熟

前述のように弱視者や一部の発達障害者を対象に特別支援教育において、タブレットコンピュータを用いた学習環境の整備が進められている。こうした試みはデジタル教科書の導入と相補的な形で今後一層の進展が予想される。そのため、タブレットコンピュータを用いることで、障害のある受験者が初等中等教育課程で利用してきた学習環境に近いユーザインターフェイスでの試験問題閲覧が可能となる。これは、受験者の受験環境に対する学習コストを低減しスムーズな解答作業を実現することに加え、受験者の誤解に起因する受験トラブルを防止する上で大きな助力となる。

1.4.1.4.3 パスワード管理によるセキュリティの確保

公平性の担保のためには不正の防止が重要である。タブレットコンピュータを用いれば、試験問題の閲覧を教科単位（センター試験の試験時間割単位）で管理するパスワード機能を実装できる。具体的には、各時間割の冒頭に試験監督者が当該時間割の教科のパスワードを告知し受験者はこれを入力して試験問題を閲覧可能とする運用が実現できる。

1.5 研究行程の概略

期間内における研究の進捗はおおむね以下の通りであった。

- (1) 5月：開発するシステムの主要要件を定式化した。
- (2) 6月～7月：(1)を踏まえ、外部有識者を交えてシステムに必要となる詳細機能を議論した。併せて、障害のある受験者への配慮に関わる全般的な意見・情報交換を行った。
- (3) 8月～9月：(2)を技術仕様に定式化する。センター試験過去問を本システムで表示できるよう電子データ化すべく、その形式・要件を決定し、製作を委託した。
- (4) 10月～12月：システム開発の発注を行い、システム開発を行った。受託者の技術的見地からの指摘を適宜聴取した。12月末に納品された。
- (5) 1月～2月：納品物の評価を行った。また、本システムを題材として、コンピュータを用いた受験配慮あるいは音声読み上げを用いた出題の特性について考察した。

2 システム要件の決定と外部有識者会議による精査

2.1 主要要件の設定

上記の認識に基づき開発するシステムの主要要件を設定した。ここにいう主要要件とは、本システムの骨子を構成する

- (1) 拡大表示の方式
- (2) 音声読み上げの方式
- (3) (1), (2)を実現するためのコンテンツフォーマット

であり、以下それぞれについて述べる。

2.1.1 拡大表示の方式について。

拡大表示は、弱視児教育で基準とされている 22 ポイントの文字サイズ相当の倍率を中心に任意の倍率に無段階調整可能とする。併せて、18 ポイント及び 26 ポイントに相当する倍率の実現に留意する。加えて、26 ポイント相当を超える高拡大倍率への対応を考慮する。受験者の閲覧作業の煩雑化を防ぐためにリフロー表示を実現する。

2.1.2 音声読み上げについて。

音声読み上げに用いる音声データは、予め収録したオーディオデータを再生する方式を用いる。

タブレットコンピュータ等に搭載される OS の音声合成機能では、与えられた文字列を解析して音声データを生成する。日本語文章を音声合成する場合、漢字の読み方や文のアクセントは OS に搭載された読み辞書データを用いて推測される。

この方式では正しい読み上げの保証はできない。特に、地歴科目の問題文に出現する人名や国語で出題される小説中の架空の人名の正しい読み上げの保証は不可能である。この点は、読みの正しさを相当程度保証できる英語音声合成 (Digital Question Paper が利用) と大きく異なる。よって、文章の製作者が意図した読みを確実に保証することはできない。厳密さを要求される試験問題の読み上げにおいて、読み間違いを許容することはできない。

本研究では、上記 OS の音声合成機能の限界に鑑み、試験実施関係者が予め音声のオーディオデータを作成・校正し試験問題コンテンツに収録する方式 (プレコーディング方式) を採用する。プレコーディング方式には

- ・ 対象文章を人間が朗読し、文字データとオーディオデータを対応付ける方法
- ・ 対象文章の文字データから音声合成機能でオーディオデータを作成し、人間が専用ソフトウェアを使って読み間違いを修正する方法

の 2 種類の方法が存在する。今回は、人的・経済的成本に優れると言われる後者の方法で作成することとした。

2.1.3 コンテンツフォーマットについて

上記の拡大表示の要件と音声読み上げの方式、具体的にはリフロー表示とプレコーディング方式の音声読み上げに対応できるコンテンツフォーマットとして EPUB3 メディアオーバーレイズを用いることとした。EPUB3 は電子書籍の国際的な標準フォーマットである。EPUB3 メディアオーバーレイズは EPUB3 規格の一部として、電子書籍への音声読み上げ収録を実現するものである。標準規格の EPUB3 を用いることで、当該規格に対応した電子書籍制作ソフトウェアで試験問題コンテンツの作成が可能となる。専用のコンテンツ制作ソフトウェアの開発を省略して、研究開発のリソースを閲覧システムの開発に集中させることができる。た

だし、一般の電子書籍では必要とされないが試験問題コンテンツでは必要とされるような機能が必要になった場合、EPUB3 規格が想定しないので対応は難しい。EPUB3 規格を越える要求の有無の確認は、本研究の課題となる。

2.1.4 研究課題の制約

本研究の置かれた状況に従い、予め以下の二つの制約を付した。

2.1.4.1 表示対象の試験問題

本システムで表示する試験問題はセンター試験過去問データとする。本来、本システムが直接対象とするのは「大学入学希望者学力評価テスト(仮称)」の試験問題であるが、本研究着手時点ではその内容や出題形式は検討段階にある。そこでサンプルコンテンツとして利用可能なセンター試験過去問を表示対象とした開発を行うことにした。センター試験は、タブレットコンピュータ上で閲覧効率を確保する上で大きな課題となる長文を含んだ大問形式の試験である。また、縦書き、数式、化学式、地図、グラフなど個別の表示課題を多く含む。これらは「大学入学希望者学力評価テスト(仮称)」においてどのような出題形式が採用される場合でも、問題化する可能性が極めて高い。よって、センター試験過去問を対象として予め解決策を探索しておくことには十分な意義がある。

2.1.4.2 提示機能開発への集中

本システムの課題は試験問題の提示方法とし、解答入力機能の開発は行わない。「文部科学省所管事業分野における障害を理由とする差別の解消の推進に関する対応指針」を始め、弱視者や一部の発達障害者への教育上の配慮として特に重視されているのは、情報の提示方法(拡大表示や音声読み上げ付与)であり、本研究が設定する喫緊の課題も試験問題提示方法の開発である。なお、これは解答記入方法や必要な試験時間の推定(試験時間延長倍率の推定)が不要であることを意味しないことを付言する。

2.2 外部有識者会議議論内容

2.1の(1)～(3)の三つの主要要件と2.1.4.1～2.1.4.2の二つの制約に従うことを大原則とし、タブレットコンピュータを用いた試験問題提示システムの詳細を検討すべく5名の外部有識者を招いた検討会議を実施した。外部有識者に期待した知見は以下の通りである。

- ・ 弱視専門家(1名): 出版・DTP関係者としてタブレットでの試験問題閲覧に関わるニーズの指摘を行う。
- ・ 発達障害専門家(1名): 障害者の高等教育進学支援、定期テスト実施支援の経験者として、タブレットでの試験問題閲覧に関わるニーズの指摘を行う。
- ・ 障害者向けデジタルコンテンツ制作ソフトの開発・同コンテンツの製作経験者(2名): タブレットでの試験問題閲覧に関わる技術的な可否の指摘を行う。
- ・ 障害者向けデジタルコンテンツ閲覧ソフトの開発・ユーザサポート経験者(1名): タブレットでの試験問題閲覧に関わる技術的な可否の指摘を行う。

以下、全4回に渡る会議の概要を示す。

2.2.1 第1回 5月26日(火)

CBT, IRTの研究動向に関わる紹介・実演をセンタースタッフがを行い、基礎的な認識の共有を行った。

主催者側から上記 3 主要要件を大原則としたシステム開発を行う方針であることとその理由、また当該システムが必ずしも一般にいう CBT, IRT を実現する試みではないことの確認を行った。

出席者から入試配慮に限定しない障害児教育におけるコンピュータ利用についての意見聴取を行った。

2.2.2 第 2 回 6 月 23 日 (火)

タブレットコンピュータを用いた試験問題提示システムによる試験実施における具体的な課題を、平成 28 年度センター試験問題を対象として、導出・列挙・整理した。

主催者側から、海外における試験配慮や試験問題を格納する電子データのフォーマットに関する動向調査の報告を行った。

2.2.3 第 3 回 7 月 23 日 (木)

第 2 回の会議で提出された課題について、対応・解決策を議論した。

本提示システム開発の範囲を超える課題、とりわけ作題上の配慮・工夫に関わる課題が存在することについて確認した。

2.2.4 第 4 回 2 月 10 日 (水)

納品されたタブレットコンピュータを用いた試験問題提示システムの機能を確認した。

論点とされた課題の解決状況を吟味した。

実システムでの閲覧を通じて、今年度想定しなかった課題についての論点整理を行った。

2.3 実装方針の決定

第 3 回までの有識者会議の議論を踏まえ、2.1 の(1)～(3)の三つの主要要件に加えて下記の詳細方針、個別課題への対応を盛り込み、システムの実装方針とした。

2.3.1 英文におけるネイティブ英語音声合成の使用

「英語」筆記の英文を日本語音声合成機能で読み上げてオーディオデータを作成することは適当ではない。英語音声合成機能を用いてネイティブイングリッシュ音声で読み上げオーディオデータを作成する。音声を頼りに英文を理解する作業にはリスニングテストの要素が混入するとも考えられる。しかしながら、それを理由として英文に音声読み上げを付与しないこと、あるいは日本語音声合成機能を用いてオーディオデータを作成することは不適切と判断した。

2.3.2 出題困難な設問の存在の確認

音声読み上げを付与すると出題できなくなる問題の存在が確認された。具体的には「国語」において出題される漢字の読みを問う問題は、その漢字が音声で読み上げられてしまう以上出題できない。

2.3.3 音声読み上げが困難な出題分野の確認と対策

音声読み上げを付与することで問う内容が変化する出題分野の存在が確認された。具体的には「国語」漢文ではリード文の漢文に音声読み上げを付与することで、作題者が想定する漢文読解作業が成立しない。他方で音声のみを頼りにリード文の理解を試みる場合、文中の漢字の意味から内容を推察することができない。

今回は試みにセンター試験点字冊子問題の過去問に準拠した問題を製作することとした。日本語点字は漢字を用いず仮名分かち書きで記述するため、漢文を表現しようとする場合、

上記の音声読み上げと同種の問題が生じる。センター試験点字冊子問題の漢文問題では出題内容を大きく変更しつつ同程度の難易度の出題が行われてきた。これに準拠する形で、同種の対応を行うこととした。

とりわけ、音声読み上げを用いた出題においては、試験問題の形式的な同質性を維持することは容易ではない。本研究ではむしろ試験問題としての実質的な同等性を確保することに注力する方針で開発を進めた。

2.3.4 表示内容のスクロール方法について

電子書籍等のコンテンツを画面上でスクロールさせながら閲覧する場合、行単位で随時表示が移動して行く連続スクロール方式と、1画面を1ページとして扱い画面単位で切り替えて行くページめくり方式の二つの方式が考えられる。図表のような通常の文字とは異なるサイズの要素が含まれる場合、連続スクロール方式ではスムーズなスクロールができない。試験問題には、しばしば図が用いられるため、ページめくり方式には一定の利点がある。

2.3.5 ユーザのニーズに従った画面表示の調整機能

表示のカスタマイズ機能が必要である。具体的には、表示文字の書体は複数のフォントから選べること、文字と背景の色が選択できることが望まれる。

2.3.6 音声読み上げの速度変更機能

読み上げの音声速度は、受験者が本人の聴覚認知特性や試験時間の制約に従い、任意に調整できるべきである。

3 開発したシステムの概要

3.1 開発するシステムの仕様と発注の概要

上記の議論の結果を仕様として整理し、「障害を有する受験者への出題に供する『任意倍率表示と音声読み上げを付与した試験問題提示システム』の開発」として発注を行った。仕様の主な要点を以下に示す。

- ・ EPUB3 に則り作成された試験問題コンテンツの表示
- ・ EPUB3 に規定されるメディアオーバーレイ機能に則り作成された音声付与試験問題コンテンツの表示・音声提示
- ・ 画面に表示される文字のフォントサイズの任意倍率への変更機能と倍率変更に伴うリフロー（文章中の折り返し位置の自動変更）機能
- ・ 背景色と表示色間のコントラストの変更機能
- ・ 画面に表示される文字のフォント種類変更機能
- ・ 音声読み上げに自動追従する画面表示機能
- ・ 試験実施に期待されるコンテンツ（試験問題）単位での閲覧制限機能
- ・ 試験問題の音声読み上げに期待される自動読み上げ停止機能
- ・ 試験問題の音声読み上げに期待される報知音出力機能

さらに、これらの仕様を実現する具体的・技術的な詳細として下記を設定した。

- ・ Windows PC, iOS, Android プラットフォームのいずれかで動作するアプリであること。
- ・ タブレットコンピュータ単体で仕様に示した諸機能が動作すること。
- ・ ユーザが、スワイプ、タップ等の動作（ジェスチャ）を用いてアプリへの指示ができること。

- ・ ユーザが、画面の視認によりアプリの操作ができること。音声出力を用いないと操作できない機能が発生しないこと。

これらシステム（タブレットコンピュータ用アプリ）本体の使用を満たすシステム開発に加え、付随する事項として受託先に以下の作業を要求した。

1. 調査をおこなう必要性から、編集集中の音声付与試験問題コンテンツの規格適合性に関する検証に参加する。
2. EPUB3 規格に規定されない試験出題に関わる本アプリ固有の機能を実現するために必要となるコンテンツ記述の独自規約等に関連した議論に参加する。

3.2 編集集中の音声付与試験問題コンテンツの表示・再生の検証

開発期間中、発注者（本研究実施担当者）と受託者の間で計 5 回の会議を持ち、3. 具体的な実装方法や周辺の状況の確認を行った。第 1 回から第 4 回は遠隔会議システムを用い、第 5 回は面談にて行った。これら会議の概要は以下の通りである。

3.2.1 第 1 回（2015 年 10 月 23 日）

仕様についての確認。仕様書に示された要件について、個別に内容の確認を行い、認識のすり合わせを行った。開発プラットフォームの候補とした 3 プラットフォーム (Windows PC, iOS, Android) の本開発における長所と短所について整理した。

受託者から EPUB3 の仕様の詳細、DAISY フォーマットについての説明を受けた。

EDUPUB など本発注に関連する諸技術や製品の動向についての情報交換を行った。

3.2.2 第 2 回（2015 年 11 月 4 日）

開発プラットフォームを iOS 特にそのタブレット環境である iPad とすることに決定した。その中でも iPad Pro を優先検証対象とすることとした。

着手した実装作業についての報告を受け、今後の作業見積もりを行った。

3.2.3 第 3 回（2015 年 11 月 17 日）

開発進捗報告を受け、EPUB3 コンテンツの表示・音声読み上げに関する基礎的な部分がおおむね完成したことを確認した。次回会議において、本システム特有の機能のうち実装方法が未決定のものについて、最終的に判断することとした。

コンテンツ仕様の検討。センター試験過去問データの EPUB3 化に用いたオーサリングソフトウェアに一部不具合があり、出力コンテンツの不整合が確認された。オーサリングソフトウェア開発者に報告するとともに、目下の作業においては手作業でコンテンツの不整合を修正することとした。コンテンツに関わる論点については、後に詳述する。

3.2.4 第 4 回（2015 年 11 月 24 日）

本システム特有の機能のうち実装方法が未決定のものについての実現方法を確定した。試験問題は、電子書籍閲覧アプリで一般に用いられる「本棚」（ブックシェルフ）という管理上のフォルダに一括登録することとした。登録された試験問題は、問題ごとに設定されたパスワードにより個別に閲覧許可を制御する。この管理方式を採用することを前提に詳細を検討した。

3.2.4.1 本棚内の試験問題の並び順

一般的な電子書籍閲覧アプリでは本棚内の並び順は登録順だが、試験問題の実施を考えると、まとめて EPUB データをデバイスにコピーする方法を採るはずである。なんらかの方法で、

本棚内の試験問題の並び順が、大学入試センター側が意図した順になるようなしくみが必要と考えられる。

解決策：

JSON (JavaScript Object Notation。ウェブ上でデータをやりとりすることを目的に開発された記述形式の一種。) の本棚情報で指定された順に、各試験問題が本棚内に表示されるようにする。

3.2.4.2 試験問題の EPUB の転送方法

試験問題の EPUB は、PC と USB ケーブルで接続し、PC 用 iOS デバイス管理ソフトウェア iTunes で転送する。アプリを起動すると、EPUB データは解凍され、本棚の中に取り込まれる。解凍後は、iTunes からはファイルの存在は見えなくなる。

3.2.4.3 データ消去の方法

本棚や試験問題の消去方法を設ける必要があるが、試験中に誤ってデータが消されないようにしないとイケない。

解決策：

本システムはあくまで研究上の実装であるが、試験実施・運用に即する方式が望ましいことには変わらない。bookshelf.json を削除した上で、アプリを起動し直すと全データが消えるしくみを実装する。bookshelf.json の削除作業のための特別な UI は設けない。iTunes から USB 接続してファイルを削除する。

3.2.4.4 閲覧履歴や個人設定の消去の方法

本のデータを消さずに、閲覧履歴のみを消去する手段を設ける必要がある。また、受験者が設定した画面表示の変更 (拡大倍率や表示フォント)、音声読み上げ速度を規定値に戻す機能も必要である。

解決策：

reset.command という名前のファイル (中身は不問) を、iTunes でアプリに転送し、アプリを起動すると閲覧履歴や個人設定が全て消去されるという仕様とする。

3.2.4.5 本棚内のグループ

地理歴史・公民、数学、理科のように、本棚の中をさらにグループ分けする必要があるかどうか。分けるとすると、本棚データ形式を工夫する必要がある。

解決策：

単体コンテンツの代わりにグループを配置し、その中に複数コンテンツ (試験問題) を収容できる仕様とする。

3.2.4.6 アプリ (本システム) の実行環境に関して

タブレットデバイス iPad シリーズを対象とし、iPhone と iPod Touch に関しては動作対象外とする。OS は iOS 9 を前提とする。

3.2.5 第 5 回 (2015 年 12 月 14 日)

受託者が現段階のアプリのデモンストレーションを行った。

試験問題のパスワード管理についての詳細を決定した。

アプリの名称に、Accessibility eXtended Examination System の略として AXES という仮称を与えることとした。

3.3 コンテンツに起因する検討事項、問題点についての整理

アプリ納品に至る期間に発注者（本研究実施担当者）と受託者の間で持った会議，並びに受託者による集中的な調査から，今回用いたセンター試験過去問 EPUB コンテンツの有する限界や問題点が確認された。以下に詳述する。

3.3.1 調査対象のデータ詳細

調査対象としたセンター試験過去問は以下の通りである。平成 27 年度センター試験：地理 A，地理 B，国語，化学基礎，化学，英語（筆記），英語（リスニング），数学 II・数学 B。

3.3.2 動作確認環境

iPad Pro (iOS 9.2) 上で動作する本システム (AXES)

3.3.3 問題点

1. 四角囲み，2 桁の縦中横の文字部分についてはハイライトされない
ハイライトされているが囲み文字が全面に表示されている状態
2. 縦書きの注記「(注 1)」，傍線部「(ア) (カタカナ)」および横書きの注記「(*)」の部分(形式)は読み上げない
3. 「下線部に誤りを含むものを」とあるが下線部がどこなのか読み上げない
4. 問の選択肢を読み上げない箇所がある
5. コードが表示されている

3.3.4 論点の総括

期待した画面表示や音声読み上げを実現するためにはアプリ(本研究の場合 AXES)の機能充実に加えて，コンテンツ制作における工夫・配慮が肝要であることが確認された。

具体的には表示対象のコンテンツが

- (1) 準拠するコンテンツフォーマットの仕様を厳守すること
- (2) (1)に加えて用いるアプリ (AXES) の特性に合わせて良好な表示を得られるよう，編集上の工夫をすること(その際，コンテンツ製作者が実アプリで製作中のコンテンツを閲覧して実地に確認するのが望ましいであろう)

が必要である。

4 有識者からのコメント

第 5 回外部有識者会議 (2015 年 12 月 14 日) の場で納品された成果物の検討・評価を行った。この会議での議論を踏まえ，3 名の外部有識者に提言の任意の提出を依頼した。その際，

- ① 今回開発したシステムについて，
 - ② 受験上の配慮について，
- を主な論点とした。

5 研究実施担当者による総括

5.1 評価

当初設定した主要 3 要件——任意倍率の拡大表示，音声読み上げ，コンテンツフォーマット——を中心に，期間中に浮上した論点も含め，個別の評価を示す。以下に示す計測結果は，アップルコンピュータ社 iPad Pro 上で本システム (AXES) を動作させ実測した結果である。

5.1.1 任意倍率の拡大表示

5.1.1.1 拡大表示, リフロー表示

拡大表示とリフロー表示を実現した。図2に平成27年度センター試験「英語」筆記第5問のリード文を標準倍率(10ポイント相当)で表示した状態(左)と207%拡大(22ポイント相当)で表示した状態(右)のスクリーンショットを示す。図3に、同第5問の設問冒頭を標準倍率(10ポイント相当)で表示した状態(左)と207%拡大(22ポイント相当)で表示した状態(右)のスクリーンショットを示す。図2と図3はいずれも縮小印刷であり、原寸大ではない。

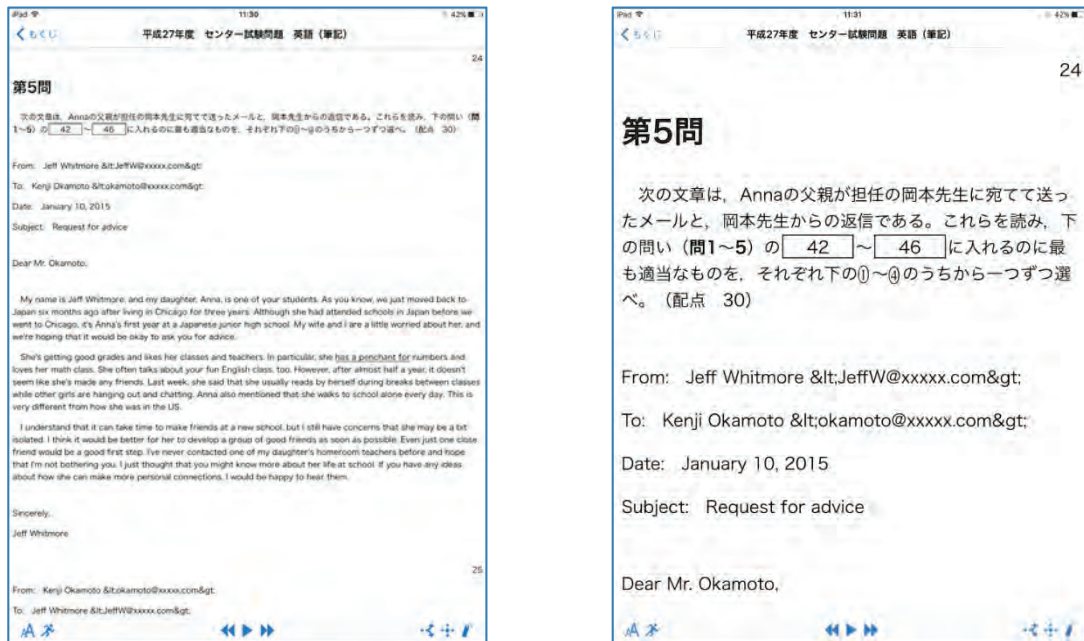


図2 標準倍率と拡大表示の比較(リード文)

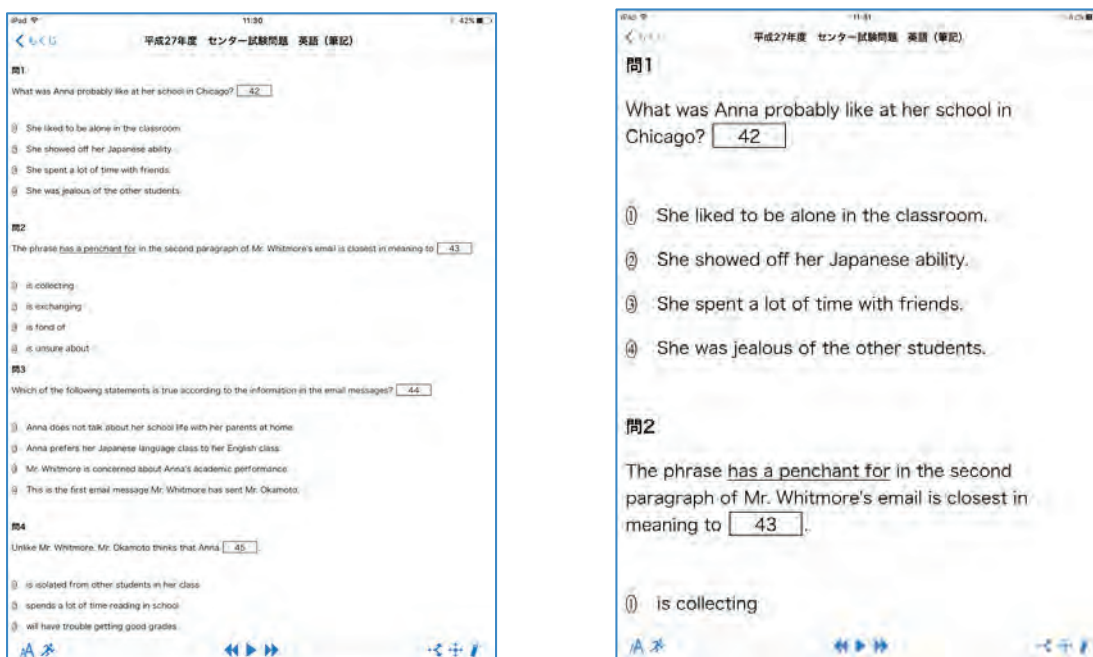


図3 標準倍率と拡大表示の比較(設問)

5.1.1.2 文字サイズ

試験問題中の主要文字種（漢字，ひらがな，数字，アルファベット大文字，小文字）について，それぞれ標準的と考えられる1文字を選び測定した。文字の高さを実測し，対応するセンター試験問題冊子中の文字との比で拡大倍率を計算した（表1）。センター試験問題冊子では10ポイントの活字を用いている。なお，測定対象とする文字やその書体によって値はある程度変化する。

表1 文字の高さの実測値と相当するフォントサイズ

計測文字	文字の箇所		紙面	AXES	AXES	AXES	AXES257%+	AXES300%+
				標準100%	拡大257%	拡大300%	アクセシビリティ機能による画面拡大1500%	アクセシビリティ機能による画面拡大1500%
「教える」の漢字「教」	国語4ページ4行目	実測した文字の高さ	3.15mm	2.92mm	7.9mm	9mm	112.91mm	136.21mm
		フォントサイズ	10pt	9.27pt	25.08pt	28.57pt	358.44pt	432.41pt
ひらがな「え」	国語4ページ4行目	実測した文字の高さ	2.9mm	2.65mm	6.5mm	7.87mm	97.47mm	104.71mm
		フォントサイズ	10pt	9.14pt	22.41pt	27.14pt	336.1pt	361.07pt
数字の「8」	数学24ページ4行目	実測した文字の高さ	2.5mm	2.13mm	5.67mm	6.6mm	87.96mm	100.62mm
		フォントサイズ	10pt	8.52pt	22.68pt	26.4pt	351.84pt	402.48pt
アルファベット小文字の「a」	外国語4ページ6行目	実測した文字の高さ	1.78mm	1.78mm	4.85mm	5.65mm	71.94mm	86.54mm
		フォントサイズ	10pt	10pt	27.25pt	31.74pt	404.16pt	486.18pt
アルファベット大文字の「D」	外国語6ページ6行目	実測した文字の高さ	2.6mm	2.49mm	6.35mm	7.68mm	97.75mm	108.66mm
		フォントサイズ	10pt	9.58pt	24.42pt	29.54pt	375.96pt	417.92pt

弱視児童生徒の利用に供するために作成される拡大教科書で用いられる文字サイズ（18ポイント，22ポイント，26ポイント）には，本システムの拡大機能のみでおおむね対応可能であることが分かった。iOSに搭載されるアクセシビリティ機能の拡大表示を併用すれば，一層の高拡大倍率の表示も可能である。拡大教科書の文字サイズでは不十分な弱視生徒が利用するとされた中心的倍率（33ポイントから36ポイント）に拡大し閲覧する場合も，利便性に重度の制約は生じないと考えられる。

5.1.2 音声読み上げ

5.1.2.1 読み間違いの回避，正しい読み上げの保証

入学者選抜試験での利用という観点から，読み間違えのない音声読み上げは最重要の課題と位置付けた。この課題は，コンテンツフォーマットとしてEPUB3メディアオーバーレイズを採用し，プレコーディング方式を実現することで解決された。試験問題を電子データ化する際に正しい読みを人為的に設定することで読み間違いは排除できる。

5.1.2.2 読み速度

今回作成した試験問題電子データの全文を音声読み上げさせた。本電子データは，EPUB3メディアオーバーレイズコンテンツ制作ソフトウェア Chatty Infty を用いて制作した。収

表2 読み上げに必要とした所要時間

教科	試験時間	標準	最速
国語	80:00	92:22	33:18
地理A	60:00	49:35	17:23
地理B	60:00	54:44	20:09
英語筆記	80:00	60:41	22:36
英語リスニング	30:00	15:03	05:48
数学2B	60:00	19:49	07:00
化学	60:00	41:09	14:46

単位(分:秒)

録した音声読み上げデータは、原則ソフトウェアの初期設定で製作しており、読み上げ速度も初期値を用いている。表 2 に、各試験問題電子データの読み上げに必要とした所要時間と当該試験の時間を示す。

全文の読み上げには長時間を要し、特に「国語」では試験時間中に読み上げが終了しないことが確認された。本システムには読み上げ速度の変更機能が実装されており、この機能を用いて読み上げ時間を短縮することは可能である。ただし、読み速度を上げた場合、技術特性上、音声に乱れが発生し聞き取りやすさは低下する。よって、標準として採用する読み速度で読み上げにどの程度の所要時間を必要とするかは重要な論点となる。

この結果は、今後

- ・ 受験配慮として音声読み上げを用いた出題を実施する際の試験時間延長倍率
- ・ 電子データに含める音声データの読み上げ速度の初期値の調整

を考える上で考慮しなくてはならない。

5.1.3 コンテンツフォーマット

前述の通り本システムでは EPUB3 メディアオーバーレイズを用いた。これによりリフロー表示とプレコーディング方式の音声収録を実現した。

5.1.3.1 縦書き表示

電子書籍の国際的な標準フォーマットは元来概して縦書き表示のための機能が不足していた。EPUB3 ではこの難点を解決しており、本システムも規格標準に則る形で縦書き表示を実現した。図 4 に平成 27 年度センター試験「国語」第 1 問の設問冒頭を標準倍率（10 ポイント相当）で表示した状態(左)と 207%拡大(22 ポイント相当)で表示した状態(右)のスクリーンショットを示す。図はいずれも縮小印刷であり、原寸大ではない。

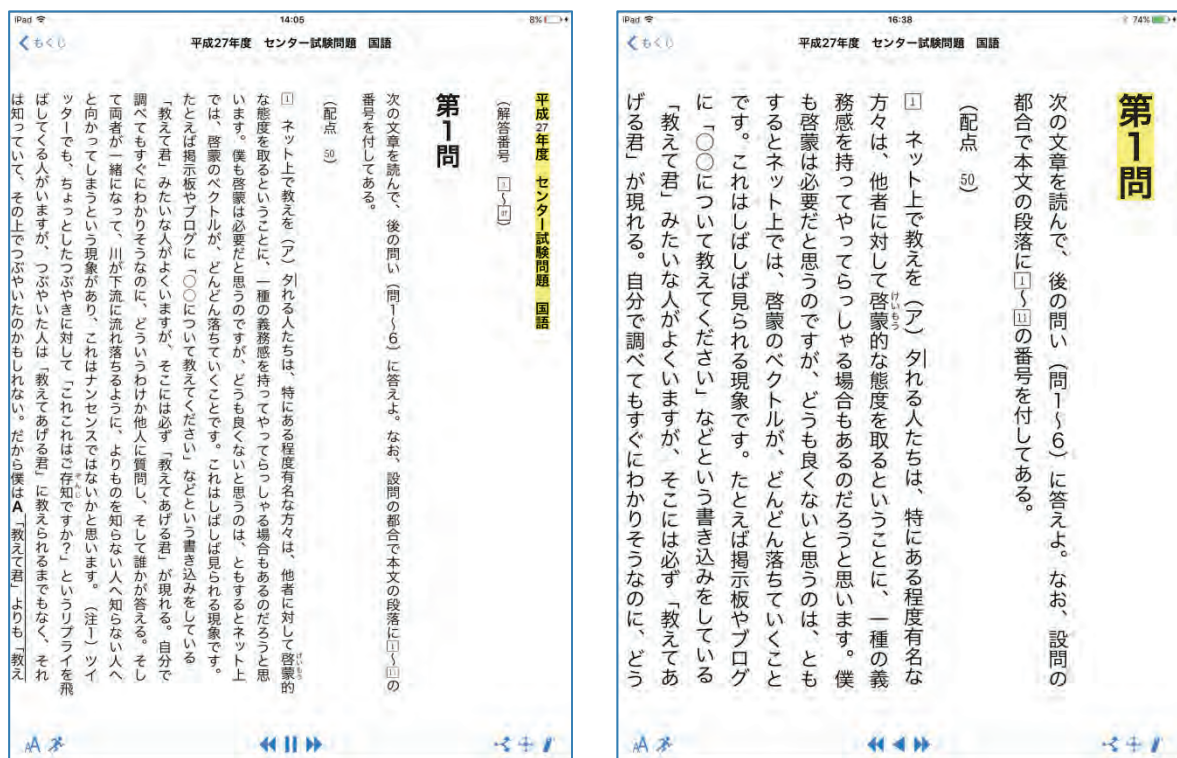


図 4 縦書き表示サンプル

5.1.3.2 リフローを考慮した問題の書き換えの必要性

今回試作した試験問題には「最も適当なものを、次ページの1~4のうちから一つ選べ」のような記載が頻出する。このような冊子を前提とした表現はリフロー型の再生環境には即していない。音声試験問題を作成する場合には問題文中の「次ページの」を「次の」のような書き換えが必要である。今後の試験問題の試作に際しては、このような音声向けの内容の書き換えのガイドラインの策定が必要である。

5.1.3.3 EPUB3 規格が想定しない機能

試験問題の提示、読み上げという個別のニーズにおいて、EPUB3 規格が想定しないが実装が望まれる電子データの要件として、読み上げの自動停止を行う箇所を指定する機能が浮上した。試験問題冊子の閲覧においては、リード文と設問の境界や大問の境界は、原則、連続的に読み進むことが想定されていない箇所である。現行の EPUB3 メディアオーバーレイズを厳守して電子データの作成、読み上げを行わせると、これらの箇所で停止することなく連続的な読み上げが行われることになる。

この問題については、EPUB3 試験問題コンテンツの再生自動停止について、自動停止したい SMIL (Synchronized Multimedia Integration Language。マルチメディアデータを表現・制御するために開発された記述形式の一種。) の audio タグに `autoPause="yes"` という属性を付けるという企画拡張を行うことで対応した。この SMIL の録音音声の再生が終わると自動停止するようにシステムを開発した。

自動停止後にハイライトする箇所をどうするかについては、今後の議論が必要である。停止した後に、自動的に次のブロックをハイライトしてから停止するのも一つの方法だが、この方法では別の問題ページの頭に飛んでしまうことが多く、受験者に解答をさせるために再生を自動的に一時停止する本来の目的から外れてしまう。今回の実装では自動停止したら停止したブロックをハイライトしたままにしている。再生ボタンを押したら同じブロックを再度再生し、今度は自動停止せずに次のブロックの再生に移動する動きとした。

解答の余裕を与える目的で自動停止した時に、どこをハイライトした状態にするのが受験者にとってベストかという課題は、今後検討する余地がある。

5.2 総括と展望

本研究では、タブレットコンピュータを用いた任意倍率による画面表示と表示内容の音声読み上げを付与した試験問題提示システムのプロトタイプの開発を行った。開発を通して、図や数式の提示と読み上げ方法、自動読み上げ停止機能の必要性などの技術的な課題が明確になった。検討を重ねた結果、技術的な課題のいくつかに関しては解決を試み、システムに試作的な方法を盛り込むことができた。それらの方法の有効性について、今後検証を重ね、テスト開始時の導入に向けて改良を迅速に進めていく必要がある。

技術的な課題とは別に、本システムの開発を通し、現在の出題方法自体の問題も明らかになった。具体的には、試験問題の表記方法が健常受験者にとっての理解のしやすさを重視しすぎたために、障害受験者への支援などを困難にしていることである。例えば、多くの教科において素材文のうち、実際に問題にかかわる部分に線と記号が付与されている。問題内容によって、「傍線+大文字アルファベット」「二重傍線+カタカナ」「波線+小文字アルファベット」などと使い分けられており、健常受験者が問題内容を取り違えないように工夫されて

いる。しかし、拡大文字が必要な受験者は、細かすぎてこのような違いを理解することが難しい。また、音声による支援を行う際、これらの情報を全て伝えようとする、複雑になり、かえってわかりにくくなってしまう。

さらに、漢文や英語の発音問題などのように、測定したい能力に読むことが密接にかかわっている問題が一部の教科で見られた。これらの問題に対してどのような支援が妥当であるかは、本チームでは判断しなかったが、重要な検討課題である。以上の課題については、障害のある受験者支援という観点だけでなく、問題作成者の作題意図なども考慮する必要がある。問題作成者などとこれらの情報を共有し、協議を重ねていくことが必要になってくるであろう。

先の中央教育審議会答申では、学力評価テストについて、表3のような提案もなされていた。「CBT方式での実施」が実現されれば、障害受験者だけでなく、全受験者がコンピュータで受験することになり、本システムの開発過程で得られた知見を十分活用することができる。その一方で、英語の「話す」試験や記述式試験など、新しい解答方法の試験が実現される場合、新たな配慮が必要になってくる。今後の議論を注視しつつ、必要な配慮の導入に向けて検討を続けていくことが求められる。

表3 学力評価テストの案と障害受験者への影響

学力評価テストの案	障害受験者への影響
知識・技能だけでなく、思考力・判断力・表現力も評価	発達障害の一部は意図が曖昧な問題への解答がより困難になる
「読む」「聞く」「書く」「話す」の四技能を総合的に評価する英語試験	「話す」に関して、聴覚障害・肢体不自由への配慮が必要
記述式を導入	肢体不自由・発達障害への配慮が必要（解答方法とともに漢字や誤脱字などのルールも必要）
CBT方式での実施	デジタル支援技術を利用しやすくなる
1点刻みの評価から段階別表示へ	公平性の考えが変化すれば、配慮自体をしやすくなる

平成 27 年度
大学入学希望者学力評価テスト（仮称）フィージビリティ検証事業 報告書
「総合報告書」

平成 28（2016）年 3 月 31 日発行

大学入試センター研究開発部（編） （研究代表者 椎名久美子）

発行所 大学入試センター研究開発部

〒153-8501 東京都目黒区駒場 2-19-23

電話(03)3468-3311(代)

印刷所 倉敷印刷株式会社
