

令和 3 年 3 月 2 4 日

独立行政法人大学入試センター

「大規模入学者選抜における CBT 活用の可能性について（報告）」の公表について

大学入試センターにおいては、大学入学共通テストへの CBT 活用の可能性について検討を行ってきましたが、検討内容を取りまとめた「大規模入学者選抜における CBT 活用の可能性について（報告）」を公表します。

（公表資料）

「大規模入学者選抜における CBT 活用の可能性について（報告）」

大規模入学者選抜における CBT 活用の可能性について (報告)

令和 3 年 3 月

独立行政法人 大学入試センター

まえがき

独立行政法人大学入試センター（以下「大学入試センター」という。）では、国内外の各種試験の CBT（Computer-based Testing）化の動向を踏まえ、10 年近く前から、研究開発部に所属する教育工学やテスト理論等の教員により CBT に関する研究を進めていたが、高大接続改革や情報教育の振興といった政府の施策の中でも、大学入学共通テスト（以下「共通テスト」という。）の CBT 化の検討が求められた。このため、研究開発部における調査研究に加え、外部委員による会議体として大学入学共通テスト企画委員会の下に CBT 活用検討部会を設置するなど、テスト理論、情報・情報技術、初等中等教育、高等教育などの各分野の専門家に御参画いただきながら、共通テストにおける CBT 活用の可能性について検討を重ねてきた。この度、これまでの検討の一つの成果として、本報告書を取りまとめるに至った。

第 1 章では、ハイスタークスの使われ方をしている大規模一斉試験として実施されている現行の共通テストの特徴を述べた上で、その試験を PBT（Paper-based Testing）で実施することに伴う課題について説明する。続く第 2 章では、共通テストにおいて CBT を活用した場合、どのような試験になるのか、現行の共通テストがどのように改善されと考えられるのかについて述べている。

一方、共通テストを CBT で実施する際にも様々な課題があり、そのことについて第 3 章、第 4 章で整理している。第 3 章では、共通テストを紙と鉛筆ではなくパソコンやネットワーク等を活用して実施する場合に生じる課題と必要な対応について、第 4 章では、共通テストを CBT-IRT で実施する場合に、パソコンやネットワーク等の活用に伴う課題に加えて生じる課題と必要な対応について説明している。

大学入試センターにおいては、今後も、本報告書で示す CBT を導入していく上での課題を踏まえ、引き続き調査研究に取り組んでいくことが重要であると考えている。本報告書は、共通テスト以外の様々な試験における CBT 活用の検討にも資する内容であり、関係各所で御活用いただければ幸甚である。

本報告書の作成に御協力くださった各位に対し、改めて心から感謝の意を表する。

令和 3 年 3 月

大学入試センター理事長

山 本 廣 基

目 次

<本 文>

はじめに	1
------	---

第1章 PBT で実施する現行の共通テスト	3
-----------------------	---

1. 大規模でハイステークスな使われ方をしている現行の共通テスト

- (1) 大学入試センターが実施する試験の概要
- (2) 志願者数とその内訳
- (3) 試験実施の流れ
- (4) 受験上の配慮

2. PBT での実施に伴う課題

【コラム①】 PBT での実施に伴う作業（問題冊子等の仕分け・輸送，答案の返送，マークシートの読取りなど）

【コラム②】 現行の共通テスト実施に要する大学教職員の負担

第2章 共通テストにおいて CBT を活用する意義	12
---------------------------	----

1. CBT の導入により実現される試験のイメージ
2. CBT の導入による共通テストの改善

第3章 パソコンやネットワーク等を活用した共通テストの実施	16
-------------------------------	----

1. パソコンやネットワーク等を活用した共通テストの主な課題と必要な対応

- (1) 試験場の設定，試験実施に関わる業務
- (2) ハードウェアの整備

【コラム③】 ネットワークと出題方式

- (3) ソフトウェアの開発
- (4) 試験実施時のトラブルへの対応

【コラム④】 『英語（リスニング）』試験における機器の不具合への対応

- (5) 本人確認・不正防止策

【コラム⑤】 自宅・高等学校など受験者が自身で選択した場所で CBT 受験する場合の本人確認・不正防止対策について

- (6) 障害等のある受験者への配慮

(7) 経費

(8) 対象科目（1科目から導入するか全面移行か）

(9) 試験日時

(10) 受験者や保護者を含む社会全体の理解と習熟

【コラム⑥】未来を担う子供たちのための大学入学者選抜と CBT の導入

2. パソコンやネットワーク等を活用した共通テストにより実現できること

(1) CBT の特徴を生かした試験問題

【コラム⑦】CBT の特徴を生かした試験問題の具体例

(2) 操作ログの取得・活用

【コラム⑧】操作ログの可能性

第4章 IRT に基づく共通テストの実施・・・・・・・・・・・・・・・・・・42

1. IRT に基づく試験

(1) IRT とは何か

【コラム⑨】2パラメタ・ロジスティックモデル

(2) 異なる試験問題に解答した受験者同士の能力の比較

【コラム⑩】IRT に基づく試験の得点の算出方法

2. IRT に基づく共通テストの主な課題と必要な対応

(1) 問題作成

【コラム⑪】問題バンクの構築・メンテナンス

【コラム⑫】医療系大学間共用試験の問題作成

(2) 試験問題の非公開

(3) 出題形式（大問か小問か）

【コラム⑬】局所独立性の仮定，一次元性の仮定

(4) 成績の表示方法

【コラム⑭】日本的試験文化

3. IRT に基づく共通テストにより実現できること

(1) 試験の年複数回実施

(2) 一人の受験者による複数回受験

【コラム⑮】複数回受験を認める場合に生じる公平性の問題への対応

(3) 受験者の能力の経時的な変化の把握

【コラム⑯】PISA 調査における得点の考え方～経年比較の観点～

まとめ	71
-----	----

1. 共通テストへの CBT や IRT の導入について
2. 今後求められる取組について

< 参考資料 >

参考資料 1	国内外における CBT の実施例	73
参考資料 2	共通テストへの CBT 導入に係る検討を求める政府決定・報告等	74
参考資料 3	大規模 CBT の実施・運営について（インドの事例）	77
参考資料 4	医療系大学間共用試験の例	78

< 付 録 >

付録 1	大規模入学者選抜における CBT 活用の可能性について（報告）（概要）	85
付録 2	「問題バンク構築に係る調査研究について～CBT-IRT での共通テスト 「情報」の問題作成に係るフィージビリティの検証～（報告）」	89

はじめに

- 従来行われてきた紙と鉛筆による試験（Paper-based Testing；PBT）が、パソコンやタブレット等のコンピュータを用いて行われる試験（Computer-based Testing；CBT）に変更されるケースが見られるようになっている（**参考資料 1**）。例えば、ETS¹が行う TOEFL[®]テストでは、日本では平成 12 年から CBT が導入されており、OECD（経済協力開発機構）が行う国際的な学力調査である PISA（Programme for International Student Assessment）も、平成 27 年からは全面的に CBT に移行された。また、文部科学省が行う全国学力・学習状況調査についても、現在、CBT 化に向けた検討が行われている。
- 大学入試センター試験（以下「センター試験」という。）は、科目によっては受験者数が 50 万人以上に上る大規模試験であり、かつ受験者や関係者にとって重大な結果をもたらすようなハイスタークスの使われ方をしている試験として実施されてきた。PBT での実施、具体的には紙の問題冊子・解答用紙（マークシート）等を使用し²、解答データはマークシートを光学式マーク読取り装置（Optical Mark Reader；OMR）で読み取って取得するという形での実施であった。令和 3 年度試験からは大学入学共通テスト（以下「共通テスト」という。）となり、当面は同様の形で実施する予定となっているが、この共通テストについても CBT 化の検討が求められている。
- まず、教育再生実行会議第四次提言「高等学校教育と大学教育との接続・大学入学者選抜の在り方について」（平成 25 年 10 月 31 日）において、その後、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）、高大接続システム改革会議「最終報告」（平成 28 年 3 月 31 日）において、「大学入学希望者学力評価テスト（仮称）」への CBT の導入や、項目反応理論（Item Response Theory；IRT，以下「IRT」という。）等に基づく複数回実施について検討が求められている。これらを踏まえ、平成 29 年 7 月 13 日に文部科学省が示した「大学入学共通テスト実施方針」では、「CBT の導入については、引き続き（独立行政法人大学入試）センターにおいて、導入に向けた調査・検証を行う。平成 29 年度については、問題素案の集積方法の検討及び集積等を行う。この成果も踏まえ、平成 36 年度以降の複数回実施の実現可能性を検討する。」とされている。

¹ ETS（Educational Testing Service）は、米国ニュージャージー州プリンストンに本部を置き、年間 180 以上の国の 9000 か所以上で延べ 5000 万人が受験するテストの開発、運営、採点を行っている世界最大の私立非営利テスト開発機関。TOEFL[®]テストや TOEIC[®] Program，GRE（大学院入学共通試験）などのテストプログラムを開発・運営している。

² 『英語（リスニング）』においては、紙の問題冊子・解答用紙に加えて IC プレーヤーも使用している。

- また、平成30年3月に高等学校学習指導要領が改訂され、「情報Ⅰ」が必修科目となり、令和4年度から年次進行で実施されることになったのを受けて、「未来投資戦略2018―「Society 5.0」 「データ駆動型社会」への変革―」（平成30年6月15日閣議決定）においては、「大学入学共通テストにおいて、平成36年度から必修科目「情報Ⅰ」などの新学習指導要領に対応した出題科目とすることについて本年度中に検討を開始し、早期に方向性を示すとともに、コンピュータ上で実施する試験（CBT）などの試験の実施方法等について検討」とされている。政府の統合イノベーション戦略推進会議の「AI戦略2019」（令和元年6月11日統合イノベーション戦略推進会議決定）や「成長戦略フォローアップ」（令和2年7月17日閣議決定）でも、「情報Ⅰ」での導入を念頭に置いた上で、共通テストへのCBTの活用が提言された。
- もっとも、これらの報告等は、それぞれが共通テストへの CBT の導入を提案している点では共通するが、意図するところは必ずしも同じではない。すなわち、文部科学省が示した「大学入学共通テスト実施方針」は、「情報Ⅰ」などの特定の科目に限らず共通テスト全体を CBT 化することが意図されていると考えられるし、複数回実施を志向していることも明らかである一方、「未来投資戦略 2018」等の各種閣議決定では、「情報Ⅰ」という科目において CBT を導入することが焦点となっており、他教科・科目への CBT の導入や複数回実施については語られていない（**参考資料2**）。
- このように、CBT の導入について検討するとしても、何を目的として CBT を導入しようとしているのかによって、具体的な制度設計も大きく変わってくる。また、CBT を導入することによる多くのメリットが想定される一方で、CBT の導入には、従来の日本の試験文化になじまない部分も多い。そのため、CBT の導入自体が目的化することがないように、その長所と短所を明確にした上で、受験者や保護者を含む社会全体の理解を得ながら検討を進めていくことが必要である。
- 独立行政法人大学入試センター（以下「大学入試センター」という。）では、従前、研究開発部に所属する教育学やテスト理論等の専門家によって、CBT に関する研究を進めてきたところである。これらの研究をはじめとした検討の成果を踏まえつつ、大規模でハイクラスな使われ方をしている試験である共通テストにおける CBT の活用について具体的に検討するため、テスト理論、情報・情報技術、初等中等教育、高等教育などの各分野の専門家を交えて検討を行ってきた。これまでの検討の成果として、この度、本報告書を取りまとめるに至ったところである。

第1章 PBT で実施する現行の共通テスト

第1章では、共通第1次学力試験（以下「共通1次」という。）及びセンター試験時代から引き継がれる、現行の共通テストの安定的な試験実施を支えている仕組みについて示し、その上で、共通テストをPBTで実施することに起因する課題について整理する。

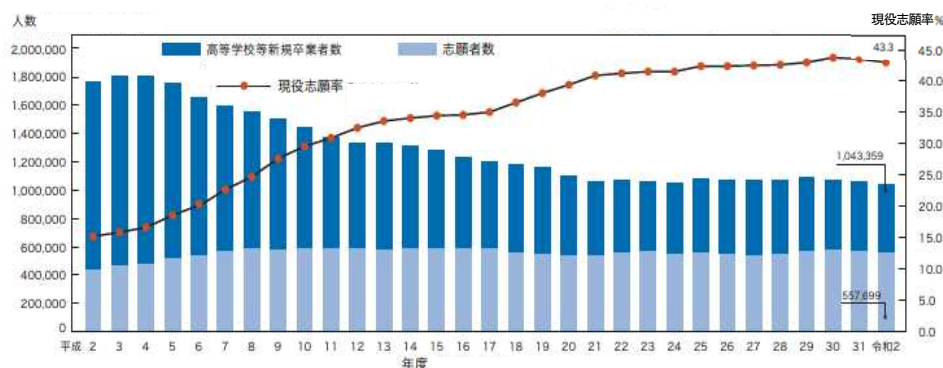
1. 大規模でハイステークスな使われ方をしている現行の共通テスト

(1) 大学入試センターが実施する試験の概要

- 大学入試センターは、昭和52年に設置されて以降、約40年にわたり、「大学に入学を志願する者に対し大学が共同して実施することとする試験」（独立行政法人大学入試センター法第3条）を実施してきた。共通1次からセンター試験を経て共通テストに至るまでの試験だが、これらは、同一期日に同一問題を用いて実施する、世界的にも類を見ない大規模試験である。
- 大学入試センターはこれらの試験に関し、「一括して処理することが適当な業務を行う」（同法第13条）とされている。これまで共同実施大学と大学入試センターは、緊密な連携体制の下で、一度も途切れることなく試験を実施してきた。こうした積み重ねもあり、これらの試験は、我が国の大学入学者選抜に寄与するとともに、大学及び高等学校の関係者からも信頼を得ていると言える。

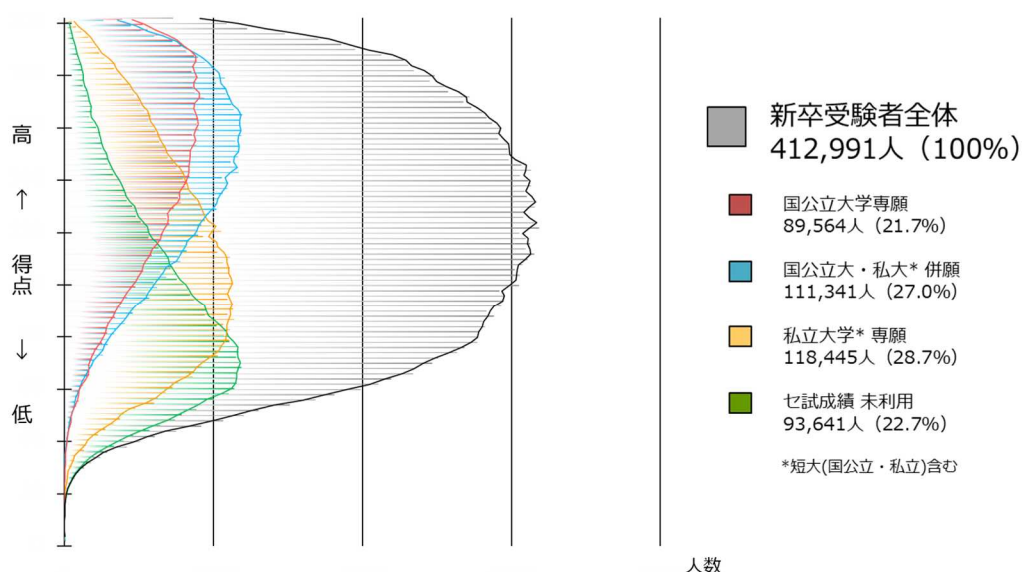
(2) 志願者数とその内訳

- 平成2年度から実施されたセンター試験では、国公立大学の志願者を対象に5教科7科目を一律に課す共通1次を改め、新たに私立大学の志願者にも対象を拡げ、大学の特色に応じて教科・科目を選択できる「アラカルト方式」を取り入れた。共同実施大学の増加に伴い、センター試験が導入された当初は約41万人だった志願者数は、近年では50万人を上回るほどに増加してきた。平成31年度センター試験からは少子化の影響もあり減少に転じているものの、令和3年度共通テストの志願者数は約53万人であった（【図1】）。



【図1】 センター試験の志願者数及び現役志願率の推移

- 志願者の内訳も多様化している。センター試験導入期には国公立大学志願者（私立大学の併願者を含む）の割合が大きかったが、近年では、私立大学専願者や、試験の成績をどの大学の出願にも利用しないいわゆる「成績未利用者」の割合が増加している。共通テストにおいても、当面は、各受験者の共通テスト受験目的は多様であると考えられる。また、この成績未利用者層は、他の受験者層に比べ、得点分布のピークが低得点域にあることが分かっている（【図2】）。このため、共通テストの受験者も、従前想定されていた受験者層から多様化していることが示唆される。



【図2】平成24年度センター試験『英語』における新卒受験者の得点分布状況

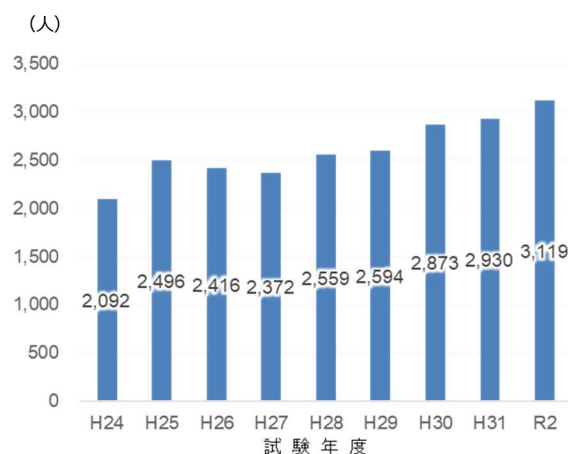
(3) 試験実施の流れ

- 試験運営に当たっては、大学入試センター職員と共同実施大学の教職員が中心となって実施上のミスやトラブルを最小限にとどめるための徹底したリスク管理を行い、公平性を確保した試験の着実かつ円滑な実施を支える仕組みをとっている。
- 試験問題については、単に知識だけではなく思考力や応用力等を問う出題上の工夫が行われており、大学及び高等学校等の関係者からは総じて完成度が高い良問であると評価されてきた。また、問題作成には約2年間をかけ、その間に問題漏洩が起こらないよう徹底した機密保持を行うとともに、約600人の大学教員等による問題作成及び複層的な点検のプロセスが採られている。なお、印刷した問題冊子等を各大学へ輸送する際にも、厳重な警備体制を敷いている（【コラム①】）。

- 試験当日には、離島を含む全国約 700 か所に設置した試験場で、延べ約 16 万人の大学の教職員が試験運営、試験監督、受付、警備等の業務に従事している。全ての試験場において円滑に試験を実施するため、大学の教職員は、大学入試センターが作成する各種マニュアルに沿って業務に当たっている。また、大学入試センターとの連絡・交信は、専用の電話・ファックス・パソコンを使用して、緊急時の対応や不明点について逐一確認することができる体制としている。
- 昨今、試験実施にかかる共同実施大学と大学入試センターの負担は、センター試験導入時と比べて大きくなっていると言える。例えば、出題教科・科目は、センター試験が導入された当初は 5 教科 18 科目だったが、現在では 6 教科 30 科目へと増加し、志願者の受験パターンも複雑化している。試験時間に関しても、9 時 30 分に開始し、終了時刻は、受験上の配慮で試験時間の延長を行う場合は 19 時 35 分に、『英語（リスニング）』に使用する機器の不具合等への対応が発生した場合は 21 時近くに及ぶこともある。こうした状況は、重大なミスやトラブルにもつながりかねないが、試験実施に関わる個々人の努力や使命感が安定的な試験実施を支えている。

(4) 受験上の配慮

- 大学入試センターでは、病気・負傷や障害等のために受験上の配慮を希望する志願者から試験前にあらかじめ申請を受け、専門家による審査を経て配慮内容を決定している。志願者専用の電話やファックスを設け、年間を通しての個別相談にも対応し、一人一人のニーズに応じたきめ細かい配慮を実施している。配慮を希望する志願者は増加傾向にあり、令和 2 年度センター試験の配慮決定者数は 3,119 人に上る（【図 3】）。



【図 3】センター試験における配慮決定者数の推移

- 主な配慮としては、
 - ① 解答方法や試験時間に関する配慮（例：点字解答，文字解答，チェック解答，代筆解答，試験時間延長）
 - ② 試験室や座席に関する配慮（例：別室設定，座席指定）

- ③持参して使用するものに関する配慮（例：拡大鏡等使用許可、補聴器装着等許可）
 ④その他の配慮（例：拡大文字問題冊子、『英語（リスニング）』の免除）
 などがある（【表1】）。

- なお、過去の例に、文字の拡大倍率や画面の照度を自由に調整できるようタブレット型パソコンを使用して受験するという配慮が、視覚区分で配慮申請を行った受験者等に対して許可されたことがある。

【表1】 配慮申請における区分及び主な配慮事項³

区 分	主な配慮事項
視覚	時間延長、点字解答、文字解答、拡大文字問題冊子、拡大鏡等使用許可
聴覚	手話通訳、監督の指示事項の文書伝達、補聴器装着等許可、『英語（リスニング）』の免除
肢体不自由	時間延長、チェック解答、代筆解答、別室設定、座席指定、車椅子使用許可
病弱	別室設定、座席指定、車椅子使用許可
発達障害	時間延長、チェック解答、拡大文字問題冊子、別室設定、座席指定
その他	別室設定、座席指定、車椅子使用許可

2. PBT での実施に伴う課題

- 上述のように、共通1次及びセンター試験は、ミスやトラブルが最小限に留まるよう試験実施上の工夫を行い、約40年にわたり実施されてきたところであり、これは共通テストの実施にも引き継がれている。このような試験実施方法は、各大学における個別入試も含め、大学入試のスタンダードとして社会的にも定着している。しかしながら、共通テストのような大規模でハイクラスな使われ方をしている試験をPBTで実施する場合には、以下のような課題を伴う。

① 出題・解答形式に制約がある

紙の問題冊子により出題する場合、紙上で表現できる問題形式でしか出題できない。解答方法としても、現行の共通テストでは多肢選択式問題の出題が主となっている。

また、受験者の解答情報として取得できるのはマークシート上に記載されたマークの結果のみであるため、受験者が解答に至るまでの思考過程などに関する情報は取得できない。

② 問題冊子・解答用紙等の印刷、輸送・保管、配付・回収が必要（【コラム①・②】）

紙媒体での実施であるため、問題冊子・解答用紙等の印刷、輸送・保管、配付・回収が必要

³ 表が示すのは各区分の主な配慮事項である。実際には、この表に示すような区分にとらわれることなく、志願者の症状や状態等に応じた配慮を行っている。

になる。現行の共通テストでは6教科30科目の試験を50万人以上の志願者に対して実施するため、問題冊子・解答用紙等の分量は膨大であり⁴、50万人分の試験問題の印刷にかかる日程を確保する必要がある。印刷が終わった問題冊子・解答用紙等は倉庫に搬入され、倉庫内で大学ごとに必要な部数に仕分けられ、輸送される。各大学において、送られてきた問題冊子・解答用紙等を、各試験場や試験室の必要な部数に仕分ける。試験当日、各試験室で試験監督者が問題冊子・解答用紙等を受験者一人一人に配付し、試験終了後に答案を回収する。

試験終了後には答案が各試験場から大学入試センターに返送される。交通状況や気象など条件が悪い場合には、予定どおりの回収が難しくなる可能性もある。大学入試センターで答案を受領した後、解答を電子化した上で採点を行うが、採点完了までには約10日間を要する。

この一連の作業は、特に事前に試験問題の内容が漏洩することがないように、全ての工程を厳重な警備体制の下で行っている。こうしたこともあって、試験実施に要する費用（令和2年度予算では約98億円）の約4割（令和2年度予算では約41億円）が問題冊子・解答用紙等の印刷・輸送等の経費⁵となっている⁶。なお、この経費は印刷や運送、警備などを担当する事業者を支払っているものであり、各大学等の教職員及び大学入試センター職員が行う作業分に相当する経費は含まれていない。

③ 試験問題は1バージョンのみで同一時刻一斉実施が必須

センター試験は、本試験、追試験ともに全国で同一時刻に一斉実施されていた。同一時刻一斉実施が求められる理由の一つとして、試験問題が本試験、追試験ともにそれぞれ1バージョンのみであり、試験実施日時が異なると、後ろの日程で受験する者に試験問題が漏洩するおそれがあることが挙げられる。

しかし、同一時刻一斉実施の場合、試験はいわゆる「一発勝負」となり、病気や事故等の事情によって受験できなくなる（あるいは、受験はできても、試験問題の内容や当日の体調等に影響される）というリスクを各受験者が抱えることになる。

さらに、新型コロナウイルス感染症などの流行性疾患の感染拡大や大規模な自然災害発生により、試験の実施自体が不可能になる⁷という社会が抱えるリスクもある。

⁴ 令和2年度センター試験では、問題冊子の印刷に約1.8億枚（B5換算）、拡大文字問題冊子の印刷に約53万枚（B4換算）、マークシートの印刷に約530万枚の紙を使用した。

⁵ 『英語（リスニング）』で使用するICプレーヤーの製造・輸送経費を含む。

⁶ ただし、CBTで実施する場合もハードウェア・ソフトウェアの整備、試験実施時のトラブルへの対応等のために相当な経費を要する（第3章1.（7）で詳述）。

⁷ 例えば、令和3年度共通テストでは、本試験1日目の令和3年1月16日に北海道留萌地方で暴風雪警報が発令されたほか、市内の公共交通機関も運休となったことから、稚内市の試験場では1日目の試験が全て中止された。また、平成7年度センター試験時には、本試験翌々日の平成7年1月17日に阪神・淡路大震災が発生し、試験終了後の答案返送に大きな影響を与えた。また、その週の土曜日、日曜日に予定されていた追試験についても、公共交通機関に甚大な被害があったため兵庫県以西の受験者が追試験場（京都府）に行けず、急きょ福岡県に追試験場を設置する措置を講じた。

なお、同一時刻一斉実施の場合は、おのずから各教科・科目の試験時間にも制約が生じる。現行の共通テストは2日間の日程で実施されているが、この日程の中で、各教科・科目の試験時間を長くすることは難しい⁸。十分な出題数を確保できないと出題に偏りが生じるなどして試験としての信頼性・妥当性の確保ができないという指摘もある中で、今後、各教科・科目の試験時間を長くする必要がある生じた場合にも、現状では対応が困難ということになる。

【コラム①】PBT での実施に伴う作業（問題冊子・解答用紙等の仕分け・輸送、答案の返送、マークシートの読取りなど）

共通1次時代は、試験を利用する大学は志願者に5教科7科目を課す国公立大学だけであったため、大学には各教科とも同じ部数の問題冊子・解答用紙等を送付することで足りた。平成2年度試験からセンター試験となり、私立大学を利用対象に加え、さらに、利用教科・科目を各大学が自由に指定できるアラカルト方式を導入した。共通テストもこの方式を踏襲しているが、この方式の下での試験実施に関する作業は膨大である。以下に、実際に行っている主な作業を紹介する。

○各試験場への送付部数の算出

アラカルト方式で実施する場合、教科（試験時間）によって志願者数が異なるため、志願者全員が全教科を受験する前提で問題冊子・解答用紙等を送付すると、教科によっては残部が大量に出ることになる。平成23年度試験までは、原則として、過去の受験者数の実績や当該年度の志願者数を考慮した上で教科ごとの必要部数を推定して送付するという方法を基本的に採っていたが、この方法では各教科の受験者数が想定を上回った場合に問題冊子・解答用紙等が足りなくなるというリスクを避けられない⁹。このため、平成24年度試験からは受験教科の事前登録制を導入し、各試験場の各教科の受験者数についてあらかじめ予測できるようにしている。

それでもなお、試験当日に不測の事態で問題冊子・解答用紙等の不足が生じることもないよう、志願者数の5%程度の数を予備分として用意している（なお、この予備分については、使用しなかった場合、大学のオープンキャンパスでの配布等、有効活用を図っている。）。

○問題冊子・解答用紙等の搬入・仕分け

試験時間割の変更や受験科目の組合せの増加に加え、共同実施大学が年々増加して試験場が増えたことにより、試験場に送付する問題冊子・解答用紙等の搬入・仕分け作業は煩雑を極め、要する時間や労力も大きくなっている。近年では、問題冊子・解答用紙等を倉庫に搬入し、試験場

⁸ 令和3年度共通テストにおいては、一般の試験時間の場合、試験終了時刻は1日目18時10分、2日目17時50分だが、受験上の配慮（試験時間延長）により1.3倍の試験時間が認められた場合は1日目18時55分、2日目18時35分、1.5倍の試験時間が認められた場合は1日目19時35分、2日目19時10分である。

⁹ 実際、平成9年度試験の「公民」において、ある試験場で問題冊子が不足し、急きょ近隣の試験場から不足分を輸送して実施するという事態が発生した。

を担当する大学別に仕分ける作業に約 30 日間を要し、一連の作業には延べ約 600 人の作業従事者を要している。

○問題冊子・解答用紙等の輸送

仕分けられた問題冊子・解答用紙等は、輸送専用コンテナに収納した後、全国各地の各大学宛てに、輸送業者の全面的な協力を得て立案した輸送計画に基づいて発送することになる。輸送計画は効率的な日程で確実に各大学に届くように練られているため、早朝から夜間まで発送し続ける日もある。また、輸送途中に問題漏洩等があるてはならないため、これらの計画は重要な秘匿事項であり、特別な警備体制を採っている。

○答案の返送

試験終了後、答案はそれぞれの試験場で枚数等の確認をした上で、大学入試センターに返送される。採点をできるだけ速やかに行うため、返送作業は発送作業よりも更に短い日程で行っており、早いものは本試験終了日の夜に到着し、3 日後までには全ての答案が大学入試センターに届くことになる。

大学入試センターに返送された答案は、再度、枚数等を確認した上で処理される。まず、輸送専用コンテナに入っている答案を取り出して専用の答案箱に移し、計数機を用いて枚数を数え、それが各大学からの送付書に記された枚数と一致するか確認する。

○マークシートの読取り、採点

試験場ごとの答案枚数の確認が終わると、光学式マーク読取り装置（OMR）を用いて答案を読み取り、そのデータを電子計算機に転送して採点を行う。読み取られる答案は、50 万人分以上、約 350 万枚に及ぶ。確実な読取りのため、この作業は装置を変えて最低でも 2 回行っており、約 4 日間で延べ約 700 万枚のマークシートを読み取ることになる。

なお、配慮が必要な受験者については、マークシートへの解答ではなく、点字や文字、チェック等による解答が認められている。こうした解答については、OMR で直接読み取ることができないため、大学入試センターの教職員等が複数人一組で解答データを作成し、内容に誤りがないことを繰り返し確認する。そして、このようにして作成した解答データについて、当日出席した受験者のマークシートが全て読み取られているかなどを厳重に確認した上で、採点を行っている。

【コラム②】現行の共通テスト実施に要する大学教職員の負担

【コラム①】では、主に大学入試センターの視点から見た PBT 実施に関する作業を紹介したが、【コラム②】では、共通テストの実施にあたる大学の教職員が、どのような作業を行っているのかを紹介しよう。一般には試験監督としての役割が目されがちだが、実際には、試験当日はもちろん、準備の段階から試験実施後まで、様々な対応を行っている。特に、大規模な PBT であるがゆえの業務としては以下のようなものがあり、これらに係る負担は CBT の導入により軽

減されることが期待される。

○試験前の問題冊子・解答用紙等の受領・仕分け・保管

各大学では、事前に大学入試センターから送られてきた問題冊子・解答用紙等を、各試験場や試験室に必要な部数に仕分ける必要がある。この事前仕分けは、担当する者の家族に受験予定者がいないかどうかの確認から始まり、試験場の規模にもよるが、ほぼ1日を要することもある。

また、仕分け終えた問題冊子・解答用紙等は、盗難や漏洩が発生しないよう試験当日まで厳重な体制で保管する必要がある。問題冊子・解答用紙等の量が膨大であれば、それに応じた規模の保管場所を確保しなければならない。

キャンパス外の試験場への輸送はもちろん、同一キャンパス内の輸送であっても、大学入試センターから大学への輸送時と同様に、輸送中のセキュリティに細心の注意を払うようにしている。

○試験前の会場設営

現行の共通テストのように同一時刻に一斉実施で行われる試験の場合、受験者が試験日の特定の時間帯に、一斉に来学することになる。このため、必要に応じて、試験場までのバス会社など公共交通機関に対して、増便依頼を行っている。

また、大学によって状況は異なるが、受験者を収容するため、多くの教室を試験室として使うことになる。このため、試験前の数日間はキャンパス全体に入場規制をかけた上で、試験室のある建物の清掃・修繕・空調及び照明音響機器の動作確認、学内はもちろん学外の動線の確保等の環境整備を行っている。試験日以前に積雪があった際には、事務職員総出で除雪作業を行うなど、万全の体制をとるようにしている。

○試験監督

試験監督者は、大学入試センターから配付される「監督要領」に基づき、確実に試験監督業務を遂行することが求められる。試験監督を担当する教員全員に対して、事前の説明会が行われるが、絶対に誤りが許されない試験監督の業務は、担当する教員にとって大きなプレッシャーであり、試験日の前夜は眠れないという声を聞くこともある。

平成24年度のセンター試験では、全国98の試験室において地理歴史、公民の問題冊子の配付誤り（「地理歴史、公民」の2科目受験者に対して地理歴史及び公民の2冊の問題冊子を配付して実施すべきところ、いずれか1教科の問題冊子しか配付しなかった等）が生じたことがある¹⁰。こうしたことが生じないよう問題冊子・解答用紙等の配付も含めて、細心の注意を払う必要がある。

問題冊子・解答用紙等の配付以外にも、正確な時刻で試験の開始・終了の合図をすることはもちろん、マニュアルに示されている数多くの受験者への指示を、所定の時刻に確実に行う必要が

¹⁰ このような配付誤りを避けるために、平成25年度試験から地理歴史、公民の両問題冊子をビニールで包装した状態で受験者に配付することになった。

ある。また、試験室内の巡回により試験監督を行うが、巡回時には靴音も極力立てない、特定の受験者への注意は口頭ではなく紙で行うなど、受験者の解答の妨げにならないよう配慮している。なお、巡回しながら受験者の出欠状況を確認して受験状況調査票（マークシート）に記録するのも、試験監督者の業務である。

○試験後の答案の返送

試験が終了し、各試験室から答案を回収した後は、複数人で枚数を確認した後、計数機も利用して念入りに全ての答案があるか確認を行い、その上で、大学入試センターへ返送する。複数の試験場を設置している大学の場合、全ての試験場からの解答用紙の到着を待つ必要があるため、その業務は夜遅くまでに及ぶ。また、万一にでも答案の紛失等あれば、受験者をはじめ甚大な影響が生じ得るため、取扱いには常に最大限の注意を払っている（なお、過去にこのような事例はない。）。

以上、PBTで行われている現状を見てきたが、CBTの導入により、問題冊子・解答用紙自体が不要になったり、試験日が分散したり、試験監督業務の一部がコンピュータで行われるようになったりすることで、ここまで列举してきた業務が軽減されるならば、それは大学関係者の負担を軽減することにつながるだろう。さらには、例えば、問題冊子・解答用紙等の配付漏れや試験時間のカウントの誤りといったヒューマン・エラーがなくなることで、受験者にとってもより安心して受験できる環境を提供することができると考える。

一方で、CBT試験についても、現時点では決して万能というわけではない。万が一にでも、受験者の解答データが失われるようなことが生じれば、取り返しがつかないことになる。また、ネットワークやCBT機器に由来するトラブルが発生した場合などに、必ずしもICTに詳しくない大学の教職員が迅速・適切に対応できるとも考えられないため、そうした実務面も含めて、確実に実施できる体制の検討が必要である。

第2章 共通テストにおいて CBT を活用する意義

1. CBT の導入により実現される試験のイメージ

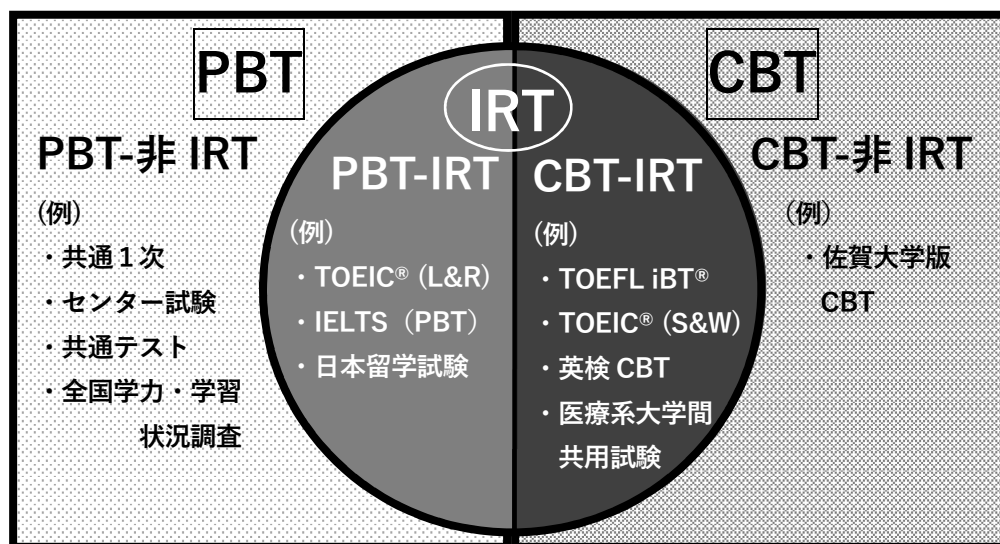
- 第1章2. に挙げた PBT 実施に伴う課題のいくつかは、共通テストを CBT で実施することにより改善される可能性がある。また、「はじめに」にもあるとおり、近年、他の試験実施機関において従来 PBT で行われてきた試験が CBT での実施に変更されるケースが見られるようになっており、共通テストについても、このような時代の潮流を踏まえ、CBT での実施について検討する必要性は高い。
- もっとも、一口に「CBT の活用」と言っても多様な形態があることには留意が必要である。例えば、「紙媒体（多肢選択式）の問題冊子・解答用紙等を、パソコン上での実施に置き換える」という、単純に試験実施のツールを紙からパソコンに変更するという活用方法も考えられるが、これだけでは CBT で実施するメリットを十分に生かせるとは言えない。CBT を導入する以上は、多肢選択式問題のみではなく「CBT の特徴を生かした試験問題を出題する」ことや、「解答に至る思考のプロセスに関する情報も併せて取得する」といった、パソコン上で出題・解答することを生かして多様なニーズに対応できるという CBT ならではのメリットを生かせるよう検討することが必要だろう。
- さらに、IRT¹¹は CBT との親和性が高い。IRT に基づく場合、統計的品質管理を行った試験問題を大量に蓄積したデータベースを構築するなど、試験問題に関する多くの情報を管理しながら試験を実施するからである。このため、「CBT を導入した上で IRT に基づく実施とし、複数回実施を可能にする」ことも考えられる。
- すなわち、第1章2. で挙げた従来の試験方式が抱えてきた課題を解決するための CBT の形態もそれぞれ異なる（【図4】）。

¹¹ IRT については、第4章で詳述。



【図4】従来のセンター試験の課題と CBT の活用方法

- 以上を踏まえると、CBT の多様な在り方を整理するための観点としては大きく二つあり、第一に、パソコンやネットワーク等を活用するかどうか（その上でパソコン上で出題・解答することを生かして多様なニーズに対応するかどうかは別途検討）、第二に、IRT に基づいた実施とすることかどうか、であると考えられる。
- これを整理すると、【図5】のように四つの類型として示すことができる。なお、本報告書では、(IRT を用いるかは保留した上で) パソコンやネットワーク等を活用することに伴う論点については第3章で、IRT に基づいた試験とすることに伴う論点については第4章で、それぞれ検討する。



【図5】CBT, PBT, IRT の関係 (イメージ)

2. CBT の導入による共通テストの改善

- 第1章2. で指摘した PBT 実施に伴う課題のうち、①の「出題・解答形式に制約がある」と

いう課題と、②の「問題冊子・解答用紙等の印刷、輸送・保管、配付・回収が必要」という課題に関しては、パソコンを用いて、試験問題や解答をデータで配信・回収することで改善される可能性がある。

① 出題・解答形式に制約がある

→ パソコン上で出題・解答することで、多様なニーズに対応できる

紙媒体ではなくパソコン上で出題・解答が行えるようになれば、以下のような可能性が広がる。

◆ 多様な方法での出題・解答が可能

マルチメディア（動画、音声等）の豊富な利用など、多様な方法での出題・解答が可能になる。例えば、総合型選抜でタブレット型パソコンによる学力検査を実施している佐賀大学では、理科の試験において、実験の動画を視聴した上で実験過程について考えたことを解答させる問題や、英語のスピーキング試験において、自らの悩みについて語るネイティブ・スピーカーの動画を視聴した上で、そのネイティブ・スピーカーへのアドバイスを英語で話して解答させる問題が出題されている（【コラム⑦】）。

◆ 操作ログの取得・活用

CBT で試験を実施する際、解答に加えて受験者の操作ログも取得できるような設計が実現できれば、受験者が解答に至るまでの過程等の情報も得ることができる（【コラム⑧】）。

② 問題冊子・解答用紙等の印刷、輸送・保管、配付・回収しての実施

→ 電子データにより試験問題・解答を配信・回収できる

共通テストをPBTで実施する場合、問題冊子・解答用紙等の印刷、輸送・管理、配付・回収などが必要だが、CBTを導入することで、電子データによる試験問題の配信や解答の回収ができるようになる。これにより、以下のようなことが可能になると考えられる。

◆ 受験者数の増減への柔軟な対応

PBTで実施する場合、全国約700か所の試験場において、全ての受験者に確実に必要な問題冊子・解答用紙等を行き渡らせるためには、出願を受け付けた後から必要部数を印刷するのでは時間的に間に合わない。そのため、受験者数を出願前に予測して印刷しなければならないが、電子データにより試験問題・解答が配信できるようになれば、受験者数の増減にも柔軟に対応できるようになると考えられる。

◆ 問題訂正等への迅速な対応

現行の共通テストでは、試験問題の内容を徹底して精査しているが、上述のように、必要な問題冊子の部数が膨大であるため、早い段階から試験問題の印刷を開始する必要がある。その一方で、どれだけ精査を重ねて問題を作成しても、例年、試験問題を印刷した後に訂正事項が発見されるケースが生じている。そうした場合には、問題訂正紙を配付する等の対応を行っているが、CBTであれば、その実施方式にもよるが、試験実施の直前の時期まで実

際に出題される問題自体を修正することが可能になるため、問題訂正紙の配付等が不要となる。

◆ **より効率的な採点の実現**

上述のように、マークシート方式で行っている共通テストのような試験の場合、全国の試験場から解答用紙を回収した上で、それらを OMR で読み取って処理を行うことが必要となる。

一方、CBT の場合には、初めから電子化された状態で解答を回収できるため、返送作業やマークシートの読取りによる解答の電子化の作業が不要となり、採点に要する時間を短縮できる。

③ **試験問題は 1 バージョンのみで同一時刻一斉実施が必須**

→ **試験問題を複数バージョン用意して複数回実施ができる**

IRT に基づいて試験を実施することで、異なる試験問題に解答した受験者同士の能力を比較することが可能となる。これにより、以下のようなことが可能になると考えられる。

◆ **試験日時の複数設定が可能**

試験日が本試験は年 1 回のみ、救済のための追試験や再試験も年 1 回のみといった場合、受験者には病気や事故等の事情により受験できなくなる（あるいは、受験はできても、試験問題の内容や当日の体調等に影響される）というリスクが常につきまとっていたが、試験日が年複数回設定されれば、後日再挑戦できるようになる。

また、新型コロナウイルス感染症などの流行性疾患の感染拡大や大規模な自然災害発生により、全国同一時刻一斉実施の試験は実施不可能に陥るおそれがあるが、複数回実施が実現すればそのような不測の事態にも対応できる。

◆ **一人の受験者による複数回受験**

従前のように一人の受験者が受験できる回数が 1 回のみだと、試験当日にたまたま低い得点しか取れなかった場合でも、それが当該受験者の能力と判断され、大学入学者選抜で用いられることになる。しかし、一人の受験者が複数回受験できるようになれば、受験者にとっては、また大学にとっても、より実力が反映された結果に基づいた選抜を行うことができるようになると思われる。

◆ **受験者の能力の経時的な変化の把握**

従来のセンター試験では、異なる年度のセンター試験の試験結果の比較は不可能だった。しかし、IRT に基づいて試験を実施することで、年度をまたいで受験者の能力を比較することが可能となれば、異なる受験者集団の能力を比較できるなど、中長期的に試験結果を比較できるようになる可能性がある。

第3章 パソコンやネットワーク等を活用した共通テストの実施

本章では、共通テストを CBT で実施する場合、どのように実施することになるのか、そのためには何を準備する必要があるのかなど、実施方法について検討した上で、パソコンやネットワーク等を活用して共通テストを実施する場合の主な課題と必要な対応について述べる。なお、本章においては、IRT の導入に関する検討については一旦保留し、IRT に関する論点については第4章で詳しく検討する。

1. パソコンやネットワーク等を活用した共通テストの実施方法と主な課題

- 本節では、パソコンやネットワーク等を活用して共通テストを実施する場合の実施方法や主な課題、必要な対応等について述べる。

(1) 試験場の設定、試験実施に関わる業務

- はじめに、CBT 試験をどこで実施するのかという試験場の問題について検討する。現在行われている CBT 試験のほとんどが、民間事業者が運営しているテストセンターを利用している。一方で、共通テストは、これまで大学等（高等学校会場を含む）を会場として行われてきており、CBT を導入した場合でも、引き続き大学等の施設を利用することも考えられる。
- まず、テストセンターについては、CBT で行われる様々な試験の試験場として、民間のテストセンター運営事業者によって運営されている。テストセンター運営事業者による直営である場合もあれば、さらに業務委託等で別会社によって運営されている場合も見られる。国内における大手テストセンター運営事業者の一つである X 社が全国で運営するテストセンターの数は、令和2年12月時点で約180か所（全都道府県、110自治体以上に設置）であるが、その多くは都市部に立地している。
- テストセンターを試験場として共通テストを実施する場合、ハードウェアはテストセンターのものを利用することになると考えられる。また、試験場の準備、試験問題の保管・管理、試験監督等の試験実施に関わる業務についても、他の CBT で行われている試験と同様に、主としてテストセンターを運営する事業者が担当することも考えられるが、これまで試験監督等を担ってきた大学の教職員の関わり方については議論が必要であるが、このことについては本項で後述する。
- 一方で、現行の共通テストの試験場である大学等に CBT 環境を整備して、CBT の試験場とすることも考えられる。共通テストの試験場となっている全ての大学に CBT 環境を整備する場合、試験場の立地や数は共通テストと同様（全国約700か所）となる。この場合、ハードウェアは大

学等のものを利用し、試験実施に関わる業務も、共通テストと同様に、主として大学等が担当することが考えられる。もっとも、CBT の場合には、ハードウェアの整備、整備後の保守、トラブル対応など含めて、実施者の負担も大きいことから、財源等を含め、実現可能性については慎重な議論が必要である¹²。

【表 2】共通テストを CBT で実施する場合の試験場

	既存の テストセンター で 実施する場合 	従来どおり 大学等 で 実施する場合 
試験場	・立地：都市部中心 ・試験場の数：約 180*	※現行の共通テストの会場を使用。 ・立地：共通テストの試験場と同様 ・試験場の数：全国約 700
ハードウェア (パソコンやネットワーク等)	・テストセンターのものを利用	・大学等で用意
試験実施に関わる業務 (試験場の準備、試験問題の 保管・管理、試験監督等)	・主としてテストセンターが担当 ※ 1 試験場当たりの座席数：10～200 程度*	・主として大学等が担当 ※ 1 試験場当たりの志願者数：最大 4,000 程度

*国内大手テストセンター運営事業者 X 社の場合

＜テストセンターで実施する場合＞

- テストセンターを利用する場合、現時点では、テストセンターを運営する事業者は全て民間の事業者であり、それぞれが、独自の基準に基づいてテストセンターを運営していることに留意する必要がある。すなわち、事業者間でテストセンターに求める要件は異なり、また、同じ事業者が運営するテストセンター間でも、パソコンの仕様やサイズ、会場の環境等が統一されているわけではない。
- また、民間事業者は既に様々な試験の実施を請け負ってテストセンターを運営していることから、新たに、合計で 50 万人以上の受験者が想定される共通テストという大規模試験を実施する上で、席数やハードウェアの整備状況、スタッフの確保などの面で対応が可能なのか、十分に検証する必要がある。とりわけ、試験を複数回実施する場合には、膨大な人数が受験する可能性がある。加えて、民間事業者には、経営状況等によって、テストセンターの縮小・閉鎖やテストセンター運営事業自体からの全面的な撤退の可能性もあることも、十分に踏まえておく必要がある。
- さらに、現時点では、国内大手テストセンター運営事業者の一つである X 社が全国で運営する

¹² 大学等にある既存のパソコンやネットワーク等を用いる等、財源の在り方も含めた検討が必要である。なお、理論上は、大学入試センターが機器等を調達して、各会場に輸送して持ち込むことも考えられるが、購入や保管、保守等のためのコストが膨大になることから現実的ではないと考えられる（(2)で詳述）。

テストセンターの総座席数は 4,000 余り（1 か所当たりの席数も 10～230 程度）と限られる¹³。仮に、受験者が 50 万人とすると、現行のような同一時刻一斉実施の形での試験を行うことはもとより、20 日間に分割して実施することとしても席数は全く足りず、別途、【表 3】にある座席数を確保することが必要になる。なお、【表 3】のシミュレーションによると、令和 3 年度共通テストの試験場が所在する全ての自治体に臨時テストセンターを設置すると仮定した場合、テストセンターを使用するための経費だけで、1 科目の場合は 1 人当たり 5,492 円、6 科目の場合は 1 人当たり 33,767 円が必要となる。ただし、この金額は、整備するハードウェアの性能や配置するスタッフ数を必要最小限に抑えた場合の額である。また、この金額には、問題作成に必要な経費やソフトウェア開発に必要な経費等は含まないため、実際に試験を実施するために必要な経費は更に高額となる。

【表 3】テストセンターで共通テストを実施する場合の経費シミュレーション

	各都道府県に 1 か所設置		令和 3 年度共通テストの試験場が所在する全ての自治体に設置	
科目数	1 科目	6 科目	1 科目	6 科目
設置座席数	5,526	33,023	5,670	33,178
延べ座席数	550,045	3,300,150	550,322	3,301,178
必要な経費	18 億 2,358 万円 (1 人当たり 3,647 円)	127 億 1,497 万円 (1 人当たり 25,430 円)	27 億 4,606 万円 (1 人当たり 5,492 円)	168 億 8,334 万円 (1 人当たり 33,767 円)

＜試算に当たっての仮定＞

- ・受験者は 50 万人。
- ・試験時間は 1 コマ 60 分。
- ・1 日当たりの試験コマ数は 5 コマ、試験を開催する日数は 20 日間。
- ・受験回数は 1 人 1 回まで（複数回受験不可）。
- ・整備するハードウェアの性能や配置するスタッフ数は必要最小限に抑える。例えば、生体認証・顔認証、身体検査、監視カメラによる監視等は行わない。
- ・問題作成に必要な経費、ソフトウェア開発に必要な経費は含まない。

※試験コマごとに異なる試験問題を使用する場合、5×20 日間=100 の問題セットが必要。

(Y 社の報告に基づいて作成)

- また、上述のように X 社も含めて、民間のテストセンターは、そのほとんどが都市部に立地しているため、離島・へき地を含む地方部に在住する受験者が不利にならないよう配慮が求められる（【コラム⑮】）。

¹³ この 4,000 余りの座席は他の CBT 試験の実施に使用されているため、仮に共通テストを X 社が運営するテストセンターにおいて実施する場合は、座席の増設や臨時会場の設置等の対応が求められる。他にも複数の大手テストセンター運営事業者があるが、状況は同様と考えられる。

- 次に、試験実施に関わる業務を民間事業者であるテストセンター職員が担うことについて検討する必要がある。共通テストは、「大学に入学を志願する者に対し大学が共同して実施することとする試験」（独立行政法人大学入試センター法第3条）としての性格から、試験実施に関わる業務（試験場の準備、試験問題の保管、試験監督等）は大学の教職員が担ってきた¹⁴。一方、現在テストセンターで行われている CBT の試験については、民間事業者であるテストセンターの職員が試験実施に関わる業務全般を担当している。テストセンターの職員は普段から様々な試験実施に関わる業務を行っており、こうした業務に熟達していて、大学の教職員の負担の軽減に寄与する可能性はあるが、大学が共同して実施する共通テストの在り方として、どのような形で民間事業者の活用があり得るのかどうかについては、国民的な議論が必要である。
- テストセンターを利用しつつ、試験実施に関わる業務を大学の教職員が担うという実施方法も考えられるが、その場合は、大学の教職員が円滑に試験実施に関わる業務を遂行できるような支援体制の整備が求められる。例えば、監督責任者は大学教職員が務めながらもテストセンターの職員も試験監督業務に協力するなど、新しい試験実施の在り方も考えていく必要があるだろう。

<大学等で実施する場合>

- 次に、テストセンターではなく大学等で CBT を活用した共通テストを実施する場合、試験実施に関わる業務は大学の教職員が担うことが想定される。しかしながら、パソコンやネットワーク等のトラブルへの対応など、技術的な専門性が求められる業務への対応も大学の教職員が行うとすれば、PBT での実施時とは異なる新たな負担が生じると考えられる。この場合、試験実施に関わる業務が円滑に実施されるよう、試験監督業務を担う大学教職員への支援（簡便な試験実施システムの構築、CBT に起因するトラブルに対応するテクニカル・スタッフの派遣等）などが必要となる。

<主な課題>

- 試験実施に関わる業務（試験場の準備、試験問題の保管、試験監督等）に技術的な専門性が求められる。
- （試験場がテストセンターの場合）座席数が限られる。また、テストセンターの多くは都市部に立地している。

<必要な対応>

- 試験実施に関わる業務が円滑に実施されるよう、試験監督業務を担う大学への支援や、民間事業者の活用について検討が必要。その際、大学が共同して実施する共通テストにおいて、どの程度まで試験監督業務を民間に委託できるのかについても検討を要する。
- テストセンターで実施する場合、分割実施が前提となる。また、離島・へき地を含む地方部

¹⁴ 高等学校を試験場として使用する場合もあるが、その場合でも、監督等の試験実施業務は大学の教職員が行っている。

在住者への配慮が必要。

(2) ハードウェアの整備

○ CBT で試験を実施する場合、出題や解答はパソコン等を用いて行うことになる。使用するパソコン等としては以下のものが考えられる。

- ・デスクトップ型パソコン：学校のパソコン教室等施設や試験専用施設での使用に適している。
電源確保が必要となる。
- ・ノート型パソコン：可搬性に優れ、臨時の試験場での使用に適している。電源確保の融通が利くが、長時間の使用ではバッテリーの管理が必要となる。
- ・タブレット型パソコン：ノート型パソコンに近く、より可搬性に優れる。電源確保の融通が利くが、長時間の使用ではバッテリーの管理が必要となる。
- ・スマートフォン：受験場所を選ばない。アンケート等の実施には適しているが、試験における使用については慎重な検討が求められる。

○ また、CBT で試験を実施する場合のネットワークやサーバについても、様々な使用方法が考えられる。ここでは、次の【表4】にある WAN 方式、LAN 方式、スタンドアローン方式の三つを挙げる。

◆ WAN 方式

試験を実施する機関（大学入試センター等）が持つメインサーバと試験場となる大学等に整備されているパソコンを WAN で接続して、直接通信を行いながら試験を実施する方式である（事前に試験場に試験問題を送る必要はない。）。

◆ LAN 方式

試験を実施する機関（大学入試センター等）が持つメインサーバから、試験場に設置されている中間サーバに事前に試験問題を送り、試験当日は中間サーバ¹⁵と受験者が使用するパソコンとの間で LAN 経由での通信を行いながら試験を実施する方式である。受験者がパソコンで解答した後、解答データを中間サーバ経由でメインサーバに送信する。

◆ スタンドアローン方式

試験を実施する機関（大学入試センター等）が用意した可搬性に優れたパソコンや外部記憶媒体（USB メモリ等）に試験実施に必要なデータをあらかじめ入れておき、試験場となる大学等に運搬して使用する。試験実施時にネットワークを使用する必要はなく、試験問題や解答データはパソコン本体や外部記憶媒体（USB メモリ等）に保存される。

¹⁵ 中間サーバとは、試験場に設置される物理的なサーバであり、メインサーバと試験場のパソコンからのみアクセスでき、試験情報や受験者の解答履歴を蓄積・管理する。また、この中間サーバでは、受験者の試験実施状況をモニタリングする機能を持つ必要があると考えられる。

【表4】CBT 試験時のハードウェアの使用イメージ

実施方式	WAN方式 メインサーバとパソコンをWANで接続して直接通信を行いながら実施。	LAN方式 事前にメインサーバから中間サーバに試験問題を送信。試験当日は中間サーバとパソコンとの間でLANで接続して通信を行いながら実施。	スタンドアローン方式 可搬性に優れたパソコンを試験会場に運搬して実施（試験時にネットワークは不要）。
試験時のハードウェア使用イメージ	試験実施機関 メインサーバ（クラウド） 試験時はWAN及びLANを使用 ルータ 各試験場 WAN LAN	試験実施機関 メインサーバ（クラウド） 試験時はLANを使用 中間サーバ 各試験場	試験実施機関 メインサーバ（クラウド） 中間サーバ 各ローカルセンター 輸送 試験時のネットワークはなし 各試験場

WAN：ワイド・エリア・ネットワーク LAN：ローカル・エリア・ネットワーク

- 大学入試センターにおいてハードウェアを整備する¹⁶場合、まず、購入やセットアップ等に必要な経費の確保が問題となる。整備した後も、ハードウェアの保管はもちろん、保守を確実にすることも必要となる。
- 例えば、パソコンについては、同一時刻一斉実施か分割実施かで必要な台数は変わってくるが、最大で数十万台に上るパソコンを保管する必要がある。精密機器であるがゆえ山積み状態での保管はできないので、保管用のキャビネット等の用意は必須である。また、パソコンが入ったキャビネット等を保管するための、温度や湿度が適切に保たれた倉庫も必要となる。さらに、単に保管することに加えて、保守も必要である。保守の具体的な内容としては、まず、試験前に全てのパソコン機器が正常に作動するかどうかの確認作業が挙げられるが、確認のためには一台一台を起動させる必要がある。また、OS（オペレーティングシステム）のアップデートやセキュリティ対策ソフトウェアのアップデート等も一台一台に対して行う必要がある。
- これらの作業を大学入試センターのみで担うのは物理的に困難であるため、試験に用いるハードウェアについては、テストセンターで実施する場合も、大学等で実施する場合も、試験場において整備されたものをを用いるのが適当である。ただし、試験場において整備されているハードウェアを利用する場合、パソコンの仕様やサイズ、ネットワークの回線速度など、試験場によって

¹⁶ 『英語（リスニング）』のIC プレーヤーのように独自開発する場合を含む（【コラム④】）。

条件が異なり、例えば、問題データを表示する処理速度に差が出たり、解答のしやすさが変わったりするおそれがある¹⁷。こうした差異が受験者の利便性に関わってくる可能性がある。

- とりわけ、大規模でハイステークスな使われ方をしている試験である共通テストにおいて、ハードウェアについての差異が生じている場合、受験者に不公平感を抱かせることになりかねない。完全に同じハードウェアではなくても、少なくとも、受験者や保護者を含む社会全体の理解が得られるような合理的な範囲での差異に収める必要がある。
- 大学等を試験場として、大学等が所有しているパソコンやネットワーク等を用いる場合には、その仕様は複数の大学間、あるいは同一の大学内であっても複数施設間で異なる場合があると考えられる。また、テストセンターで実施する場合も、各テストセンター運営事業者が、試験で使用するハードウェアについて独自の基準を持っていたりするものの、同一のテストセンター運営事業者であっても、試験場によって多少の差異が生じているのが現状である。もちろん、一定の基準を示した上で新たに購入（あるいはレンタル）すれば、ハードウェアを一定程度統一することは可能だが、現実的ではない。
- なお、パソコンに関しては BYOD (Bring Your Own Device；私用デバイスの活用) とすることも考えられる。実際、大学や高等学校等の授業などでは導入されているケースもあるが、大学入学選抜での利用となると、不正行為の防止が BYOD 以外の方法に比べて難しいことや、試験中のトラブルへの対応が難しいといった課題がある。現状では、自宅受験が認められている場合を除くと、CBT で行われている試験において BYOD が認められているケースは基本的にない¹⁸。

<主な課題>

- 大規模なハードウェアの整備、及びその後の保管や保守が求められる。
- テストセンターや大学等に既に整備されているハードウェアを活用するのが合理的だが、その場合、パソコンの仕様やサイズ、ネットワークの回線速度などの条件が試験場によって異なることになる。

<必要な対応>

- 受験者や保護者を含む社会全体の理解が得られる形でハードウェアを整備するとともに、その後の保管や保守を確実に行うことが必要。

¹⁷ 特に画面をスクロールする必要が生じる長文や図表の読解問題、画面上での作業を必要とするプログラミング問題などについては、画面サイズの違いが解答のしやすさに関わってくる可能性がある。

¹⁸ ただし、2次元 CAD 利用技術者試験 (CAD (Computer-aided Design) を利用するための知識を持ち、更に図面を正しく理解して CAD を利用した作図を効率的にこなすことができる技能を証明できる試験制度) の 1 級試験が、受験者に CAD ソフトをインストールしたノート型パソコンを持参させて実施しているように、BYOD を前提として実施している CBT 試験も存在する。

【コラム③】ネットワークと出題方式

IRTには様々な実施方式があるが、受験者が正答するとより難しい問題を、誤答するとより易しい問題を出题するというように、受験者の解答履歴から逐次的に能力を測定し、その能力に応じて試験問題を出题する方式として、アダプティブ方式(コンピュータ適応型テスト(Computer-adaptive Testing；CAT))¹⁹がある。

CBTはIRTとの親和性が高いが、IRTの中でも特にアダプティブ方式で実施する場合は、パソコンとネットワークを活用する必要性が高い。これは、アダプティブ方式で実施する場合、受験者が解答する問題セットを試験開始前に固定することができず、受験者の解答履歴を踏まえて次に出题する問題を決定することになるからである。このように出题される試験問題が流動的で、試験の進捗に合わせて出题する試験問題を決定し、受験者に示すという方法をPBTで採用することは難しく、CBTの特徴を生かすことができるだろう。

なお、大規模試験の場合には、多くの試験問題を用意する必要があるが、受験者が解答する問題以外の試験問題も含めて、中間サーバや受験者用のパソコン(小型の外部記憶媒体(USBメモリ等)を含む。)に事前に配信することは、セキュリティや容量の問題から難しい。そのため、ネットワークを使用する方式の方がなじむと考えられる²⁰。

このように、CBTを活用してIRTに基づく試験を実施するとしても、出题方式によってネットワークなどのハードウェアの整備を考えていくことが必要である。

(3) ソフトウェアの開発

- CBTで試験を実施する場合、試験に関わる一連のプロセスをコンピュータ上で実行するためのアプリケーションやブラウザ等のソフトウェアが必要となる。民間事業者が開発したソフトウェアや、(テストセンターで実施する場合には)テストセンター運営事業者により開発されている各種のソフトウェアを活用することも考えられるが、現行の共通テストの性格を維持したままCBTで実施する場合、こうしたソフトウェアには、少なくとも以下のような機能が求められる。

① 共通テストの出题形式に対応する機能

現行の共通テストの形式を維持する場合、ソフトウェアには大問形式(一つのストーリーに沿って複数の小問を組み合わせたまとまり)での出题²¹のための機能が求められる可能性がある。また、新たに出題教科とすることが検討されている「情報」に関しても、プログラミング問題等のCBTの特徴を生かした試験問題を出题する場合は、そのための機能が必要となる。

¹⁹ アダプティブ方式については、第4章1.(1)で詳述。

²⁰ PISA2018もアダプティブ方式で実施されたが、45程度の問題セットで実施することが可能だったため、外部記憶媒体(USBメモリ)を用いたスタンドアローン方式で実施された。

²¹ 出题の形式(大問か小問か)については、第4章2.(3)にあるとおり、IRTに基づく実施という観点からも検討する必要がある。

② 本人確認・不正防止のための機能²²

本人確認・不正防止を厳密に行うためには、ソフトウェアにもそうした機能を実装することが求められる。

本人確認のための機能としては、例えば、事前に登録した受験者の顔写真とパソコンの前に着席している人物が同一人物かを顔認証の技術を用いて確認する機能をソフトウェアに実装することが考えられる。

不正防止のための機能としては、例えば、試験開始直前に発表されるパスコードを入力しないとテストが始められない機能、検索エンジン、電卓、漢字予測変換など試験実施時に作動すべきでないアプリケーションやブラウザは起動できないようにする機能などをソフトウェアに実装することが考えられる。さらに、パソコンに搭載されているカメラを併用して、顔の動きから不正行為の疑いのある受験者を検知する機能を備えることなども考えられる。

③ 障害等がある受験者への配慮のための機能

現行の共通テストにおいては、障害等がある受験者に対してきめ細かい配慮を行っている²³が、CBTの導入後も同様の配慮を行うためには、そのための機能をソフトウェアにおいても実装することが求められる。

文字サイズの変更機能、フォントの種類の変更機能、図表の拡大機能、読み上げ機能など民間事業者が開発するソフトウェアにも実装されている機能に加え、点字ディスプレイ²⁴のサポート、マウス操作（ドラッグ&ドロップなど）による解答が困難な受験者が物理キーボードでの解答を可能とする設計など、配慮が必要な受験者のユーザビリティも考慮したソフトウェアの開発が必要である²⁵。

- 共通テストで各教科・科目をCBTで実施する際には、上述のように、独自性の高い機能が実装されたソフトウェアが必要となるため、これまで大学入試センターで行ってきた調査研究の成果を生かしながら、共通テスト本番で利用できるソフトウェアを、場合によっては独自開発によ

²² 本人確認・不正防止策全般については(5)で詳述。ソフトウェアに機能を実装する以外にも、指紋認証の技術を用いるなどの方法が考えられる。

²³ 受験上の配慮については、第1章1.(4)で詳述。

²⁴ 点字を表示する電気機械式デバイスであり、平坦な表面に穿たれた穴からドットが上がってくることによって点字を表す機能を備えていることが一般的である。

²⁵ 大学入試センター研究開発部においても、タブレット型パソコンを用いた障害等がある受験者への配慮環境の開発・評価・整備に関する研究を進めている。開発中のアプリケーションでは、試験問題の音声読み上げ、文字や図の拡大、フォントや画面配色の変更等が可能だが、実用化に向けては、①紙の拡大文字問題冊子の作成に高度な専門性が求められるのと同様に、配慮を必要とする受験者用の試験問題の電子化にも高度な専門性が求められる、②大学入学者選抜試験という用途を踏まえた正確な音声読み上げを実現する技術を採用する必要がある、など課題がある。

り用意することが必要となる。なお、独自開発したソフトウェアについては、開発後もクラウド環境の維持、改修、セキュリティ対応等を継続的に行っていくことが必要になることも考慮に入れておく必要がある。

<主な課題>

- 現行の PBT で実施されている共通テストの特徴を生かしたまま CBT を導入する場合、独自性の高い機能を備えたソフトウェアが求められる。

<必要な対応>

- 求められる機能が実装されたソフトウェアを、場合によっては独自開発により用意することが必要（独自開発したソフトウェアを使用する場合、開発後のクラウド環境の維持、改修、セキュリティ対応等も必要。）。

(4) 試験実施時のトラブルへの対応

- CBT を実施する上で留意しなければならないのが、試験実施中にハードウェアやソフトウェアに不具合が発生するリスクを伴うことである。CBT で実施する場合、パソコンやネットワークの不具合などのトラブルを皆無にすることは困難であり、一定のトラブル発生リスクを伴うこと、また、トラブル発生の場合に備えておくことが必要である。
- 例えば、インドで約 86 万人の受験者に対して 4 日間の日程で実施されたある大規模な CBT 試験における、天災・システムトラブルへの措置を見てみよう。この試験では、まず、バックアップのデータセンターが設置されている。これにより、メインのデータセンターが利用不可能になったとしても、試験を継続することができる。また、振替試験日を設定することで、万が一、予定されたテストセンターでの試験実施が困難となった場合、試験日程の変更や追加コストは生じるが、試験中止という事態を回避できるようにしている（**参考資料 3**）。
- また、国内のテストセンター運営事業者である X 社においては、システム障害時に以下のような対応をとることとしている。

【表5】システム障害時の対応例

試験開始後 経過時間	試験場	X 社	試験を実施する機関
0 分	【障害発生の第一報】 X 社へ報告	・事象の確認 ・復旧対応	—
	【受験者への案内】 障害の発生や復旧対応を行っていることなどの状況を説明。障害復旧まで待機を案内。	【試験場への対応指示】 ・復旧対応 ・受験者対応方法の指示 【障害規模の判別】 全体障害か個別障害か判断	
30 分～	受験者から試験日の振替希望がある場合はそれに応じる。	【試験実施団体へ報告】 ・30 分以上の試験の中断及び開始遅延が見込まれる場合、影響範囲を報告。 ・試験日の振替対象者が確定後、振替対象者のリストを提供。 【受験者への連絡】 ・試験日の振替が発生した場合、個人申込みによる受験者へ振替を案内。 ・団体申込の場合は、申し込みをした団体に振替を案内。	・X 社からの第一報で、障害状況の確認・把握 ・障害規模、影響範囲等を公表

(X 社の資料を一部大学入試センターで加工)

- もっとも、こうしたトラブルの発生は、ハードウェアやソフトウェアだけに起因するものではなく、さらに、外部からのサイバー攻撃やウイルス等による被害も想定する必要がある。実際、官公庁や企業、大学、病院等において、毎年のように様々な被害が生じているように、完璧なセキュリティの構築は困難なのが実情である。とりわけ、共通テストのように注目度の高い試験の場合には、こうした攻撃の対象になる可能性も高いと考えられるが、それによって試験の実施が困難になったり、受験者の解答データが滅失・欠損したりするようなリスクについても留意する必要がある。
- 以上のような状況を踏まえると、トラブルなく試験を実施するためには、事前の対策、具体的には、試験で使用するハードウェアの保守を徹底することが極めて重要である²⁶。テストセンターで実施する場合は、各テストセンターの運営事業者がハードウェアの保守を行うことを含めた契約をすることになると考えられる。大学等で実施する場合には、原則としてハードウェアを所有している大学等が保守を行うことになると考えられるが、共同実施である大学にのみ過大な負担が集中することがないよう、慎重な検討が必要である。
- また、予想外のトラブルに対応するため、予備のパソコンを用意したり、受験者の解答データを随時サーバにバックアップしたり、あるいはサーバやネットワークのトラブルに技術的に対応

²⁶ ハードウェアの保守については、(2)で詳述。

できるテクニカル・スタッフを試験場に配置したりすることなどが必要になる。

- しかしながら、上記のような事前の対策を徹底したとしても、なお、トラブルを皆無にすることは困難と言わざるを得ず(上述のように、外部からの攻撃の可能性も想定される)、結果的に、試験を最後まで受けられなかったり、解答データの滅失・欠損が生じたりする可能性がある。特に、入試に関しては一切の実施上のトラブルは許されないと考えられていることから、トラブルが起きた場合に受ける社会からの批判や賠償責任などを懸念して、CBT の実施を引き受けるテストセンター運営事業者や共同実施大学が現れないケースが想定されることにも、十分留意が必要である。

<主な課題>

- ハードウェアの保守を徹底するとともに、トラブル発生に備えて、予備のパソコンの用意や解答データのバックアップ、各試験場へのテクニカル・スタッフの配置等が求められる。
- それでもなおトラブルを皆無にすることが困難であることについて、受験者や保護者を含む社会全体の理解を得ることも重要。

<必要な対応>

- トラブルの影響を除去できない場合、振替試験日の確保、代替の試験問題の準備等が必要。

【コラム④】『英語（リスニング）』試験における機器の不具合への対応

センター試験では、平成 18 年度試験から『英語』にリスニング試験が導入された。受験者には問題冊子・解答用紙等と IC プレーヤーが配付され、各自が個別に音源を再生して解答する。試験の実施方法については、一旦は、放送設備が整備された高等学校を試験場としてスピーカー方式による実施案が公表されたこともあったが、どの試験場でも同一の条件で受験を可能にするために検討を重ねた結果、個別音源による方式によることと決定し、現在に至るまで毎年実施してきたところである。

試験で使用する IC プレーヤーは、センター試験専用のものを独自に開発し、リスニング試験導入以降も改良を重ねてきた。IC プレーヤーを独自開発とした理由は大きく二つある。第一に、機器の不具合によるトラブルのリスクを可能な限り回避するために、機能を最小限に絞り、音声メモリーに記録された音源の再生に特化した機器としていることである。第二に、受験者の操作誤り等を極力低減するために、操作ボタンは試験に必要な三つに絞る、ボタンの色をそれぞれ変えた上で操作順を示す番号を付すなど、受験者にとって操作方法が単純で分かりやすいものになるよう工夫をしていることである。

試験当日、各受験者は指示に従い操作準備を行った上で解答を始める。解答が続行できないような機器の不具合や不慮の事故等が発生した場合は、試験を中断し、対象の受験者のみ、全体の試験時間終了後に、別の機器を使って中断した設問から解答する「再開テスト」を行う。

試験導入年には、約 50 万人の受験者のうち 400 人余りが再開テストを受験することになり、

社会的に大きな関心事となったが、それ以降、仕様や運用面の改善を重ね、近年の再開テスト受験者は 100 人程度となっている。再開テストとなる理由についても、詳細に検証した結果では、受験者の操作に起因するものが大半である。ただし、上述のような機能の絞り込みを行い、製造工程における品質管理を徹底しても、なお、機器自体の不具合が、毎年 0.002% 程度（50 万台中 10 台程度）確認されているところである。

一方、CBT の場合には、試験で使うパソコンやタブレットは、リスニング試験で使用している IC プレーヤーよりもはるかに複雑な機器である。加えて、ネットワークの安定性も必要になることから、試験に際してより様々なトラブルが生じる可能性が高い。CBT を実施する上では、こうした事例も参考にしながら検討していくことが重要だろう。

(5) 本人確認・不正防止策

- パソコンを用いて実施する試験については、PBT での実施時には想定されなかった新しい形の不正が行われる可能性がある。そのため、(3)で言及したソフトウェアへの本人確認・不正防止機能の実装も含め、新たな本人確認・不正防止策を検討する必要がある。
- 例えば、試験でデスクトップ型パソコンを使用する場合、机上の解答用紙に解答を記入する場合とは異なり、他の受験者の解答画面を容易に覗くことが可能になる。また、試験システムの画面構成が一般に知られるようになると、解答画面を見ずとも、手元の動作だけでその受験者の解答内容が把握されてしまう状況も生じる可能性がある。そのため、パーテーションを設置したり、パソコンのディスプレイにのぞき見防止機能を施したりするなど他の受験者の解答画面が見えないようにする工夫が求められる。
- 一方、パーテーションの設置などを行った場合、従来の試験室内の巡視を十分に行うなどの方法のみでは不正行為の未然防止が難しくなる。このため、生体認証（顔認証や指紋認証等）を活用した本人確認や、天井等に設置する監視カメラあるいはパソコンに搭載されているカメラを使用した監視などを導入することが考えられる。
- 試験場にこのような設備を導入する際に経費を要することはもちろんだが、こうした対応をとることについては、個人情報やプライバシー等の保護の関係からも、受験者や保護者を含む社会全体の理解を得た上で実施していくことが必要である。

<主な課題>

- PBT での実施時には想定されなかった新しい形の不正が行われる可能性があり、生体認証（顔認証や指紋認証等）を活用した本人確認や、監視カメラ等の使用など、従来とは異なる新たな不正防止策が求められる。

＜必要な対応＞

- 個人情報やプライバシー等の保護の関係から、受験者や保護者を含む社会全体の理解を得ることが必要。

【コラム⑤】自宅・高校など受験者が自身で選択した場所で CBT 受験する場合の本人確認・不正防止対策について

CBT を導入する方式の一つとして、受験者自身が、自宅や高校など受験可能な場所を選択し、自分が所有しているパソコンや、高校等に備えられているパソコンを、試験実施機関のサーバと接続して試験を行う方式（以下「自宅等における受験」という。）も考えられる。このような方式で試験を実施する場合、例えば、離島やへき地の受験者の移動や宿泊等に伴う負担を軽減することなどが期待されるが、一方では、本人確認・不正防止をどのように行うかが極めて重要となる。特に、共通テストのように、ハイステークスな使われ方をしている試験であれば、その重要性は一層高い。

ここでは、本人確認・不正防止のための手段のうち、遠隔からの試験監督及びソフトウェアの搭載機能についてまとめる。

○遠隔からの試験監督

自宅等における受験を認める場合には、遠隔地にいるスタッフがウェブカメラを経由して試験監督を行ったり、あるいは AI を活用したりすると考えられる。令和 2 年に新型コロナウイルス感染症が世界的に流行して以降、このような方式で実施される CBT 試験の普及が加速し、パソコンに搭載されたウェブカメラ等を使って、遠隔地から人間が監督したり、AI による顔認識機能等を活用したり、これら両方を組み合わせたりして試験監督が行われている。

AI の機能開発も進んでおり、事前に提示する顔写真付きの身分証明書、出願時に提出する顔写真、受験中に映る顔の三つが一致しているかを自動的に判定することも可能になっている。また、受験中の受験者の顔の向き、目線、パソコンのマイクから聞こえる音などから不正行為の可能性を検知する機能も開発されている。

ただし、現時点では、不正行為に係る最終的な指摘や判断は人間が行う必要があると考えられ、AI を活用する場合であっても、人間による監督を組み合わせることが現実的である²⁷。また、遠隔からの試験監督により試験を行う場合と、従来どおり人間による試験監督の下で試験を行う場合とを比較した、不正行為の傾向や検知件数の違い等のデータはまだ十分には得られていない。

²⁷ 新型コロナウイルス感染症の拡大を受け、自宅での受験を可能とする仕組みが整備されたテストもある（TOEFL iBT®テスト、ケンブリッジ英語検定 4 技能 CBT リングスキル等）。TOEFL iBT® Home Edition は、試験監督者にカメラを介して監視された状態で試験が実施される。受験者のみで他の誰も入っていない部屋で受験する、テスト前にカメラを動かして試験監督者に机の上及びテストを受験する部屋全体（360 度）を見せる必要がある、試験問題の漏洩防止の観点から通常の紙にメモを取ることは不可（メモは消去可能なホワイトボード等にとる）などのルールがある。ケンブリッジ英語検定 4 技能 CBT リングスキルは、AI 技術と人間による遠隔監視の両方によって実施されている。

い。遠隔からの試験監督を行うことの適否を検討する際は、十分なデータを踏まえた上で、慎重に判断することが求められる。

○ソフトウェアの搭載機能

自宅等における受験では、基本的に、各受験者が自身のパソコンを使用することが想定される。その場合、不正防止策として、試験を実施するためのソフトウェアに以下のような機能を搭載することが考えられる。

- ・画面操作の記録
- ・試験開始時に他アプリケーションを強制終了
- ・試験中の試験アプリケーション・試験画面以外の利用を禁止
- ・スクリーンショット操作を停止
- ・リモートデスクトップ機能を停止
- ・仮想環境の利用を禁止
- ・マルチディスプレイの利用を禁止

ただし、自宅等における受験がまだ普及していない中で、試験を実施するためのソフトウェアにどのような不正防止の機能が求められるかを検討するためには、さらなる調査研究や先行事例の収集が必要である。

(6) 障害等のある受験者への配慮

- 共通テストに CBT を導入する際には、障害等がある受験者のアクセシビリティを考慮した CBT 環境を整備することが求められる。例えば、文字サイズの拡大、配色・コントラスト変更、リフロー表示（文字を折り返しながら拡大表示）など画面表示の調整、音声読み上げ等が実現される可能性があるが、実現されれば、視覚障害のある受験者の受験を支援することができる。また、点字ディスプレイを使用できる環境を開発すれば、点字による教育を受けている者の受験を支援することができる²⁸。
- 一方、症状や状態等によっては CBT での受験が困難なことも考えられるが、そのような場合には引き続き PBT で実施するなど、細やかな対応をとっていくことが必要である。また、2.(1)で詳述するが、CBT の特徴を生かした試験問題（例：動画を見て解答させる問題、画面に表示された図表や文章における該当箇所を選ばせるという方法で解答させる問題、等）を出題した場合、障害等のある受験者が従来以上に解答に困難を覚えるおそれもある。したがって、共通テストに CBT を導入する際にはこうした点にも十分に留意する必要がある。

²⁸ 現行の共通テストの点字問題冊子を作成する際、図表、画像を使用しているなどそのまま点訳することが難しい試験問題については、文章中心にするなど代替問題を作成し、それを点訳している。

<主な課題>

- CBT で実施することで障害等のある受験者の受験を支援できる可能性がある一方、障害等のある受験者の中には CBT での受験が困難な者もいる。

<必要な対応>

- 可能な範囲において、障害等がある受験者のアクセシビリティを確保した CBT 環境を整備する。
- CBT での試験の受験が難しい場合については、引き続き PBT で実施するなどの対応をとる必要がある。

(7) 経費

- 共通テストをパソコンやネットワーク等を活用して実施しようとした場合、上記の(1)から(6)までで述べた以下の諸点に対応するための環境整備が必要となる。

- ・試験場及び試験実施体制の整備
- ・ハードウェア（パソコンやネットワーク等）の整備
- ・ソフトウェア（出題、解答のためのアプリケーション等）の開発
- ・試験実施時のトラブルへの対応（予備のパソコン、テクニカル・スタッフ等）
- ・不正防止策（生体認証のための機器、監視カメラ、監視体制等）
- ・障害等のある受験者への配慮（アクセシビリティを考慮した CBT 環境の整備等）

しかし、これらの環境整備のためには相当な経費を要することは当然である。

- テストセンターや大学等のハードウェアを利用したとしても、パソコンやネットワーク等を活用する場合、試験実施に要する経費が従来と比べて高額になることが見込まれている。仮に、国からの財政支援が得られない場合には、その費用を検定料等に転嫁することが必要になるが、受験者や保護者等の負担も考慮した慎重な検討が必要になる。

<主な課題>

- 試験実施に必要な経費が PBT での実施に比べて高額。
 - ・ハードウェア（パソコンやネットワーク等）の整備
 - ・ソフトウェア（出題、解答のためのアプリケーション等）の開発
 - ・試験実施時のトラブルへの対応（予備のパソコン、テクニカル・スタッフ等）
 - ・本人確認・不正防止策（生体認証のための機器、監視カメラ、監視体制等）
 - ・障害等のある受験者への配慮（アクセシビリティを確保した CBT 環境の整備等）

<必要な対応>

- 実施経費の増加に伴う財政負担について検討することが必要。

(8) 対象科目（1科目から導入するか全面移行か）

- 仮に、共通テストの一部の科目のみを CBT で実施し、他科目を PBT で実施する場合、それらを CBT で実施するために要する経費と、他科目を PBT で実施する経費（問題冊子・解答用紙等の印刷、輸送・保管等）の両方が必要となる。そのため、試験実施においては、PBT あるいは CBT のいずれかの形で実施するよりも、必要なコストは大幅に増加すると考えられる。このため、CBT に移行する場合には、全ての教科・科目について一斉に移行することが、コスト面からは望ましい。仮に、一部の科目について先行的に CBT を導入する場合には、速やかに CBT 対象科目を拡大する目途を立てるべきであり、こうした場合には、国の財政支援がない限り、実施経費を検定料等に転嫁することなどが必要となる。

<主な課題>

- 一部の科目のみを CBT で実施するのはコストパフォーマンスが悪い。

<必要な対応>

- （ほぼ）全教科・科目を CBT で実施する目途が立ってから、共通テストへの CBT の導入を開始することが妥当。
- 仮に、一部の科目のみに CBT を導入する場合は、国の財政支援がない限り、実施経費を検定料に転嫁することなどが必要となる。

(9) 試験日時

- 試験の実施日時については、同一時刻に一斉実施とするか、試験日時を複数設定して分割実施とするかについて検討する必要がある。この点、テストセンターで実施する場合には、(1)で述べたように席数が限られていることから、基本的に分割実施とせざるを得ない。大学等で実施する場合には、理論上は同一時刻一斉実施により行うことも可能だが、現状では各大学が保有するパソコンの仕様も多様であり、均質なハードウェアを全国的に整備するためには経費が膨大になる。また、トラブルへの対応を考慮しても、同一時刻一斉実施ではなく分割実施（試験日時を複数設定）の方が実施しやすいことなどから、分割実施を検討することが現実的である。
- ただし、共通テストのようなハイステークスな使われ方をしている試験において、同じ試験問題を異なる日時に実施される回で使用することは、試験問題の内容が事前に漏れることになるため、事実上不可能である。したがって、試験日時を複数設定して分割実施とする場合、少なくとも設定された回数分の試験問題を用意する必要がある。
- 異なる試験問題に解答した受験者同士の能力を比較できるようにするための手法としては、IRT を用いることが考えられる。仮に、共通テストを分割実施によって実施する場合には、IRT

を活用することが有用と考えられるが、その場合にも様々な課題を伴うことについては留意が必要である（第4章で検討）²⁹。

<主な課題>

- 受験環境の整備やトラブルへの対応を考慮すると、同一時刻一斉実施ではなく分割実施（試験日時を複数設定）の方が実施しやすい。

<必要な対応>

- 分割実施の場合、試験問題を複数パターン準備することが必要（第4章で検討）。

(10)受験者や保護者を含む社会全体の理解と習熟

- 上述のように、パソコンやネットワーク等を活用して共通テストを実施する場合、マルチメディア（動画、音声等）の豊富な利用など、多様な方法での出題・解答が可能になり、PBTでは測定が難しかった能力を測定できること等が期待されるが、一方では、試験実施にかかる経費がこれまで以上にかかると考えられること、また、トラブルが発生して試験を最後まで受けられなかったり、解答データの滅失・欠損が生じたりする可能性があるリスクがあることなどについて、受験者や保護者を含む社会全体の理解を得ることが必要である。
- また、CBTが受験者等に受容されるためには、受験者が、パソコン等の操作も含めて、CBTという試験の形態に慣れた状態にあるかどうかも重要な要素となる。このためには、初等中等教育段階でのCBTの活用も進むことが望ましいが、共通テストに関しても、解答に必要な操作を円滑に行うことができるよう、共通テストで使用するソフトウェアの操作を体験できる練習用のウェブサイトが大学入試センターが公開するなど、受験者が本番以前にソフトウェアの操作に慣れる機会を提供することが望ましい。

<主な課題>

- 共通テストへのCBTの導入は受験者にとってメリットもある一方、トラブルの発生や検定料の引き上げの可能性といったデメリットも想定される。

- CBTが受験者等に受容されるためには、受験者がCBT試験に慣れておくことが望ましい。

<必要な対応>

- CBTのメリット・デメリットについて周知し、受験者や保護者を含む社会全体の理解を得ることが必要。

- 練習用のウェブサイトを用意するなど、受験者がCBTに慣れる機会を提供することが重要。

²⁹ 共通テストをIRTに基づいて実施する際の主な課題については、第4章2. で詳述。

【コラム⑥】未来を担う子供たちのための大学入学者選抜と CBT の導入

大学入学共通テスト企画委員会 CBT 活用検討部会においては、議論の対象を明確化するため、「大学入学者選抜を目的とする大規模でハイクラスな使われ方をしている共通テストへの CBT の導入」に焦点を当てて検討が行われた。しかし、共通テストをはじめとする大学入学者選抜への CBT の導入を検討する際には、PBT を CBT で実施すべきか否かという形式面に留まらず、「未来を担う子供たちのための大学入学者選抜はどうあるべきか」という大局的な視点が本来は求められる。

試験の内容については、CBT の導入により、大学教育を受ける際に必要だが、従来の大学入学者選抜では測ることができていなかった資質・能力を測ることができるかを検討する必要があるだろう。例えば、令和 3 年 1 月から始まった共通テストは思考力等を問う問題が充実したと言われているが、資質・能力の三つの柱の一つである「学びに向かう力、人間性等」を PBT で問うのは未だ難しい。しかし、CBT を導入し、学力試験の結果と電子調査書での主体性評価とをうまく連携させれば、大学教育を受けるのに必要な力をより効果的に測定できるかもしれない。

また、試験運営についても、受験者が身に付けた力を発揮できるよう、受験者目線での設計が重要である。例えば、試験で使用するパソコンは、公平性や不正防止の観点から課題があるのは承知の上だが、個人が所有するパソコンを試験場に持参させて実施する方式について検討のそ上に載せる必要はあるだろう。これまでの入試では、合否を左右する極度の緊張感、精神的なプレッシャーにおいて発揮される受験者のパフォーマンスが測られてきた。受験者が普段使用しているパソコンを用いることは、大学入学後に求められる受験者の日々のパフォーマンスを見とることにつながるだろう。また、複数回実施について検討する際も、高等学校や大学という機関の都合だけではなく、受験者が自分で受験したいタイミングで受験できるというメリットについて十分に踏まえることも必要であろう。

以上のことを実現するためには、共通テストの性格や位置付け自体が変わる必要があるのかもしれない。質の高い受験環境をそもそも求めない、受験環境に不具合が生じた場合は再度受験できる、試験問題は第 2 学年修了程度の問題とするなど、共通テストの在り方も含めて検討するべきであろう。戦後最大の教育改革と言われる新学習指導要領実施において、学校における学びが大きく変化している。CBT を導入した新たな共通テストによって、未来を担う子どもたちに必要な学びがさらに進展していくこと、また、各方面での議論が集約された CBT の導入を含めた大学入学者選抜改革が、時代を生き抜く子どもたちを養う仕組みとなっていることを期待したい。今後も共通テストへの CBT の導入については検討が続くと思うが、CBT の導入自体が目的とされることなく、「未来を担う子供たち」のための議論が展開されることを切に願っている。

(執筆：高井潤（埼玉県立狭山工業高等学校主幹教諭）)

2. パソコンやネットワーク等を活用した共通テストにより実現できること

- 1. では、パソコンやネットワーク等を活用して共通テストを実施する場合の実施方法や主な課題、必要な対応等について検討した内容をまとめたが、その中では、実現に向けて克服すべき課題についても多く指摘した。
- 本節では、仮にそれらの課題が克服されて共通テストをパソコンやネットワーク等を活用して実施する運びとなった際に、実現可能になることを整理する。

(1) CBT の特徴を生かした試験問題

- 第2章1. でも述べたとおり、CBT を導入する以上は、PBT での試験を単にパソコン上に移すだけではなく、CBT の特徴を生かした試験問題を出题することが期待される（【コラム⑦】）。
- 共通テストにおいて CBT の特徴を生かした試験問題を活用する意義としては、これからの時代に必要な資質・能力を、より適切に測定できるようになることが挙げられる。平成30年に改訂された高等学校学習指導要領では、生徒に求められる資質・能力の育成を目指しているが、その資質・能力には、教科等横断で育成を目指す「情報活用能力」も含まれる。CBT を活用することで、文章や図表のみならず、多様なメディアから情報を取り出し、分析する力を問うといったことも可能になるだろう。
- 一方、CBT の特徴を生かした試験問題を作成する上では、これまで以上に時間や労力を要するものもあると考えられる。そうした場合、継続的・安定的に問題作成をすることのフィージビリティ（実施可能性）についても検討が必要である。
- また、障害等のある受験者への配慮を考えた場合、CBT の特徴を生かした試験問題がマイナスに働くおそれもある。すなわち、動画を見て解答する試験問題、画面に表示された図表や文章における該当箇所を選択する解答方法などについて、試験問題が CBT の特徴を生かした複雑なものになりすぎると、障害等のある受験者の解答が、従来以上に難しくなる可能性があることも考慮する必要がある。
- このように、CBT の特徴を生かした試験問題にはメリットもある一方、課題もあるため、例えば、CBT の特徴を生かした試験問題の出题は全体の中の一部にするなど、実現可能な活用方法を検討した上で取り入れることが適切である。なお、CBT の特徴を生かした試験問題を出题したことによる効果の検証も必要であるが、短期間で結論を得ることは難しいため、中長期的に継続して検証しなければならない。

【コラム⑦】 CBT の特徴を生かした試験問題の具体例

CBT を導入し、パソコンで試験を実施する場合、以下のような特徴をもつ問題を出題できる。

○試験問題の多様性

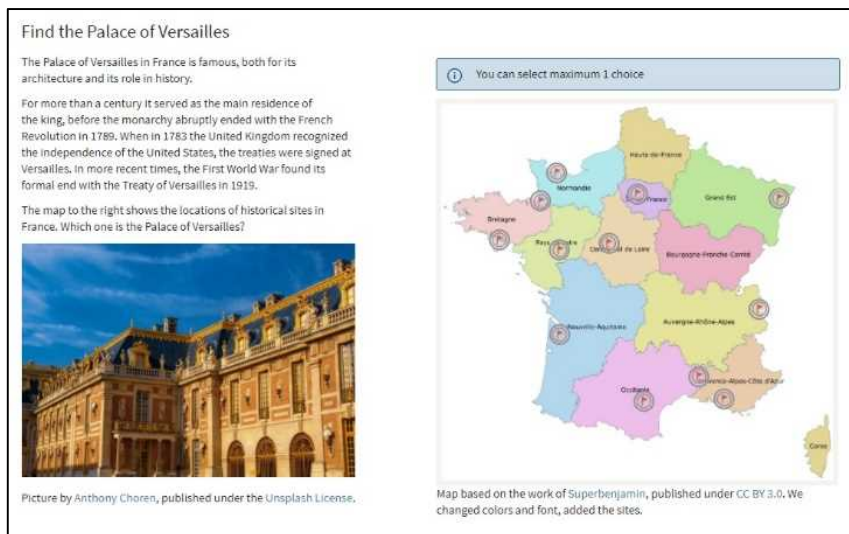
例えば、動画を見たり音声を聞いたりして解答する試験問題など、メディアを豊富に利用した問題を出題できる。



【図コラム⑦－１】 動画を用いた試験問題の例
(実験を行う動画を見た上で、実験の結果について解答する。)

○解答方法の多様性

画面に表示された図表の該当箇所を選択する、画面に表示された文章の該当箇所に下線を引く、文章を入力するなど、多様な形式の解答を要求する問題を出題できる。



【図コラム⑦－２】 図表の該当箇所を選択させる試験問題
(写真の建物が存在する都市がどこなのか、右の地図のマークをクリックして解答する。)

○受験者と試験問題の双方向性

受験者が一度送信した解答に対してフィードバックを行い、そのフィードバックを踏まえて受験者が更に試行錯誤をして解答を導くというような問題を出題できる。



【図コラム⑦-3】プログラミング問題

特に、【図コラム⑦-3】のような「情報」のプログラミングの問題については、CBTで試験を実施するメリットが大きいと言える。すなわち、プログラミング問題をPBT（マーク式）で出題する場合、問題作成者がつくったプログラムを読み解くことが中心の出題となる傾向が多い。一方、CBTで出題する場合、受験者が自分でプログラムを組み立て実行し、実行結果の確認及びプログラムの修正の過程を経ながら目標に近づけていくといった解き方をさせることが可能になる。また、複数の考え方を容認する実際のプログラミングに近い形での出題が可能になる。

<出典>

【図コラム⑦-1】佐賀大学版 CBT 「総合問題 sample」

<https://www.youtube.com/watch?v=QjQwaqq4go&feature=youtu.be>

【図コラム⑦-2】The TAO User Guide

<https://www.taotesting.com/user-guide/users/interactions/what-is-an-interaction/>

【図コラム⑦-3】プログラミング問題

本報告書付録2

(2) 操作ログの取得・活用

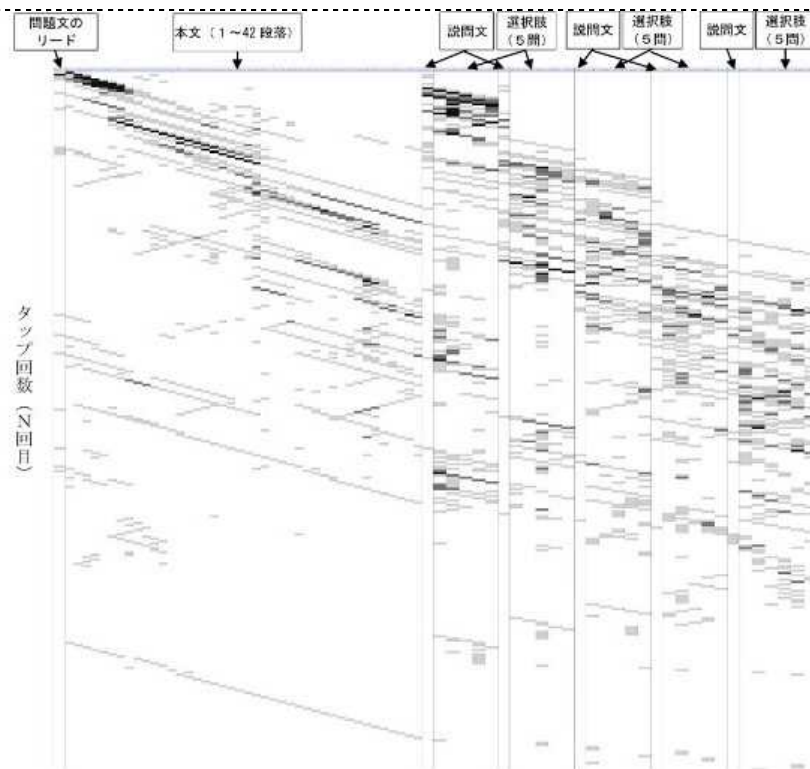
- 第2章1. で述べたとおり、CBTで試験を実施する際に、受験者の操作ログを取得して、受験者が試験問題に取り組む過程(解答プロセス)を把握できるような設計とすることも考えられる。条件が整えば、取得した操作ログを大学入学者選抜だけではなく、大学入学後の教育への活用や

試験問題の検証, 受験者の解答行為(例えば, 解答に要した時間や解答時の画面の閲覧状況など)の分析等に活用することも考えられる。

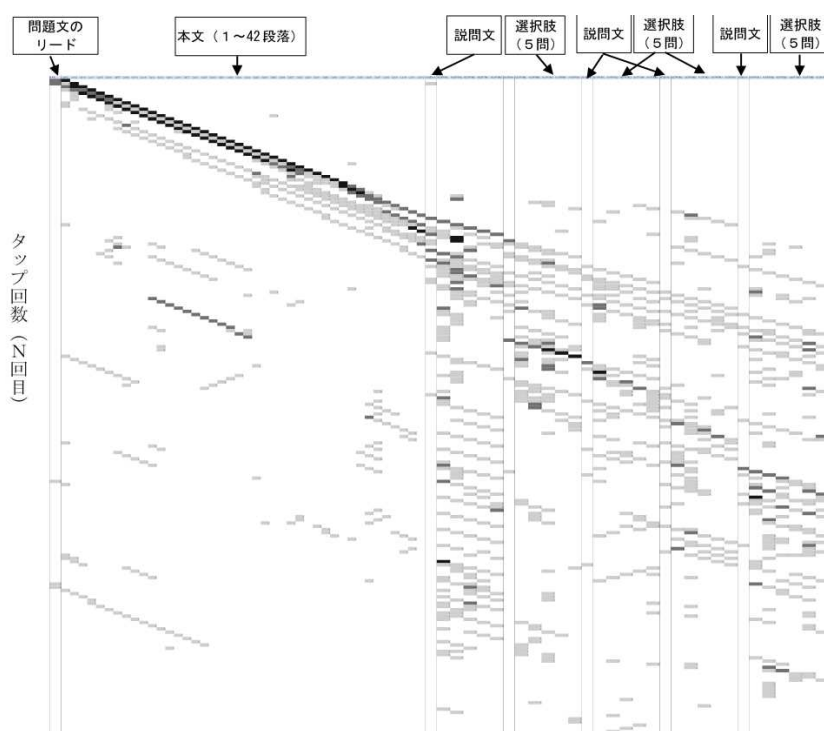
- 一方, 選抜のための試験と解答プロセスを把握するための試験問題では, 一般的にテスト・デザイン自体が異なる。このため, 仮に共通テストにおいて, 大学入学者選抜のために必要なデータに加えて, 解答プロセス等のデータも取得しようとしても, 解答プロセス等の分析が十分にできないおそれがあることに留意しなければならない。
- このため, 操作ログの取得・活用は, 共通テストへの CBT の導入の際の必須事項とすることは適当ではないが, 諸条件が整った場合には操作ログを取得し, 検証等の目的で活用することから開始することが望ましい。

【コラム⑧】 操作ログの可能性

「操作ログ」とは, コンピュータなどデジタル機器を使いながら問題解決をするときに(プログラムに規定された範囲で)自動的に残る記録のことである。CBT で操作ログを取得すれば, 解答結果だけでは分からない解決プロセスが見えてくる。もちろん, PBT でもビデオカメラなどを使えば記録は残せるが, 大人数を対象に自動処理が容易になる点で CBT はメリットがある。それによって, 例えば, 多肢選択式問題で各設問の解答をどのように選択したかというプロセスが一人一人について把握でき, それを集積した全体のパターンも把握できる。【図コラム⑧-1】及び【図コラム⑧-2】は, センター試験の現代文の問題を CBT で解いた実験参加者が本文と設問をどのような順番で読み解決しているかをヒートマップ状に表したものである。【図コラム⑧-1】の参加者は本文を読んで設問に関わる傍線部まで来ると, 設問に飛び, また本文に戻るといった解き方をしているのに対し, 【図コラム⑧-2】の参加者は本文を全て読んでから設問を順に解いている。



【図コラム⑧－１】 解答プロセスの可視化の例（本文傍線部まで→設問文→解答→本文, n=16）



【図コラム⑧－２】 解決プロセスの可視化の例（本文→設問文→解答, n=11）

プロセスが把握できると、同じ解答をした受験者でもどちらの読み方をした受験者に志願してほしいかを大学側が判断したり、傍線部周辺の段落しか読まない読み方では正答し難い問題

を出題したり、本文を読み終わるまで設問を表示しないというコントロールをかけるなど、測定したい資質・能力が発揮される認知活動をも踏まえた精緻なテスト・デザインと評価が可能になる。

そこで、次の研究では、国立大学第二次学力試験国語（評論）問題を題材に、本文や設問を出題されたとおりに再現した CBT（以下「従来版」、ただし縦書きを横書きに変更した）と、「本文全体の論旨を踏まえた上で」解答することが求められた 4 問目の設問と本文だけを同じにして、他の設問等に改変を加えた CBT（以下「改変版」）を用意し、後者でより積極的な読解プロセスが引き出せるかの比較実験を行った。改変版では、従来版の四つの設問のうち、問 1～3 の代わりに、①まず本文全体を読んで（傍線部ごとに本文を断片的に読むことを防ぐため、本文を読まない問題に進めないようにした）、②本文に含まれる三つの具体例を三つの構成要素で書き出す計 9 マスの表を手掛かりに従って埋め（【図コラム⑧－3】の下方 A～D 欄が表を埋めた例）、③構成要素にラベルを付けて（【図コラム⑧－3】の表の一番上の行）、④ラベルを詳しく説明する（【図コラム⑧－3】のラベルの下の方）という活動を用意した。その上で、従来版の問 4 を解いてみることにした。加えて、ステップの③以降は本文が読めないようにして、作った図表（メモ）を基に自分なりの言葉で解答を書くことを促した。本文の解体・再構成を CBT で促進することを狙ったわけである。

中堅私立大学生 40 人と国立大学受験者も多い進学校の高等学校 3 年生 39 人に従来版と改変版 CBT に取り組んでもらい、統合的な問 4 の課題解決成績を比較したところ、進学校生では改訂版、中堅大学生では従来版の成績がよいという交互作用が見られた。中堅大学生の従来版では本文からの書き写しが多いことで却って記述採点の部分点を得て成績が上がったのに対し、進学校生は従来版でも自分なりの言葉で解答を構成しようとして成績を落としていた。これに対し、改変版の CBT では両者とも自分なりの解答を構成しようとしていた。それが進学校生で機能し、中堅大学生で機能しなかったのは、改変版の①～③のステップで記入欄の正答率が進学校生で 62.7%、中堅大学生で 37.8%と、ステップが本文の解体・再構成の足場掛けとして働いた程度が違ったためだと思われる。これは進学校生の読解活動の積極性を示唆している。

以上の結果からは、従来版のテストだとあまり深く文章を理解せずに本文の言葉を繋ぎ合わせて解答してしまう中堅大学生でも改変版だと本文の解体・再構成をしようとする中堅大学生の従来版と進学校生の改変版の統合成績結果の点数だけでは違いが見えず（これは採点方法にもよるが）、プロセスを見ると両極の実態が見えることが分かる。このように CBT には、これまで「点数」としてしか処理されてこなかったテストの解答の背後にある受験者一人一人の読解や思考などの認知プロセスを推察できるデータを入手できる利点がある。

残り時間 16:09

※ これ以降は本文を参照せずに解答する問題です。

問 本文について、下の表に整理したい。

4. 本文中に「歴史的出来事」の存在は『理論的』あるいは『物語り』の存在であり、フィクションといった誤解をあらかじめ防止しておくならば、それを『物語りの存在』と呼ぶこともできます」とあったが、「歴史的出来事」の存在はなぜ「物語りの存在」といえるのか、本文全体の論旨を踏まえた上で、100字以上120字以内で説明せよ（句読点も1字と数える）。

[文字数0]

(1) 理論的存在	(2) 間接的証拠	(3) 理論的手続き
(1)の説明 直接的には観察できないが、間接的証拠と理論的手続きによって存在が証明されているもの。		(3)の説明 理論的存在と間接的証拠とを結びつけて、理論的存在の真実性を証明するために必要や理論や調査などの手続き。
素粒子	霜箱や泡箱、サイクロトンによって捕らえられた痕跡	A 物理学理論
前九年の役	B 衣川の古戦場	C 『物語り』のネットワーク
D 赤道や日付変更線	六分儀等の計器による計測	地理学や天文学の理論

【図コラム⑧－3】改変版 CBT 画面例（問4 を解く際）

<参考文献>

北澤武, 白水始 (2020)「CBT による多肢選択式問題の解決プロセスの解明—大学入試センター試験問題の国語既出問題を活用して」, 大学入試研究ジャーナル, 30, 44-51.

白水始 (2021)「積極的読み」を引き出す CBT 読解問題の開発 科研費基盤研究 (S) シンポジウム「e テスティング最前線」

(URL: <http://www.ai.lab.uec.ac.jp/symposium21/Shirouzu.pdf>)

(執筆：白水始 (国立教育政策研究所初等中等教育研究部総括研究官))

第4章 IRTに基づく共通テストの実施

現在 CBT で行われている大規模試験の多くは、IRT (Item Response Theory, 項目反応理論) というテスト理論を活用して実施されることが多い³⁰。本章では、IRT とは何か、共通テストを IRT に基づいて実施する場合、どのような試験になるのか、従来の同一時刻一斉実施の試験と比べてどのような違いがあるのかなどについて述べた上で、IRT に基づいて試験を実施する場合の主な課題と必要な対応について述べる。なお、IRT 自体は必ずしも CBT で行う必要はなく、実際に PBT で行っているケースもある。

1. IRT に基づく試験

(1) IRT とは何か

○ IRT とは、各受験者の試験問題に対する正答・誤答に基づいて、試験問題の特性と受験者の能力を分けて推定する統計理論の一つである。IRT に基づいて試験を実施する場合、以下のようなことが実現できる。

◆ **異なる試験問題に解答した受験者同士の能力が比較できる（そのため試験の複数回実施が可能）**

受験者が異なる試験問題に解答した場合でも、不公平性を排除して得点を算出することが可能になるため、異なる試験問題に解答した者同士の得点も比較できる。このため、例えば、IRT に基づいて共通テストを実施すれば、現行では本試験、追試験をそれぞれ年 1 回のみ実施しているところ³¹、年に複数回の試験を実施できるようになる可能性がある。

◆ **統計的品質が管理された試験問題を出題できる³²**

上記のように異なる試験問題に解答した受験者同士の能力を比較可能にするためには、難易度推定のための事前の予備調査³³を実施して調査参加者にあらかじめ試験問題に解答しても

³⁰ CBT において IRT が活用されるケースが多い理由としては、一度に受け入れ可能な受験者数に限りがあり、同一時刻一斉実施ではなく IRT に基づき複数回に分割実施をせざるを得ないという消極的な理由から、CBT であれば IRT に基づく試験を効率的・効果的に実施することが可能であるという積極的な理由まで、様々なものが考えられる。

³¹ 令和 3 年度共通テストは、新型コロナウイルス感染症の影響に伴う学業の遅れに対応できる選択肢を確保するため、本試験を 2 回実施した。

³² ただし、IRT の出題方式は多様で、その中でも予備調査による事前の統計的品質管理が必須でない方式も存在する（分冊方式やアンカーテスト方式など）。そのような方式で実施する場合は、試験問題の品質が事前に明らかにされるわけではない。

³³ 予備調査の方法としては、受験者集団に近い集団を対象にした「プレテスト」の実施や、本番の試験において新作の試験問題（案）をダミーとして出題することなどがある。予備調査については、【コラム⑪】で詳述。

らい、その解答データから試験問題の品質を推定しておく必要がある。このプロセスの中で、品質が基準を満たさない試験問題（例えば、難しすぎる又は易しすぎる試験問題、能力の高い受験者と低い受験者で正答率があまり変わらない試験問題など）を除くことになる。すなわち、試験問題の統計的品質が一定程度確保された試験問題によって、試験を実施できるということになる。

- IRT に基づく場合には、本番の試験前に予備調査を通して試験問題の統計的品質管理を行うことが一般的であるが、試験問題の品質を示す指標は「項目パラメタ」と呼ばれる。IRT モデルには、日本における多くの大規模試験で用いられている 2 パラメタ・ロジスティックモデルや 1 パラメタ・ロジスティックモデルをはじめ様々なものが存在する³⁴が、項目パラメタのうち代表的なものが以下の二つである。

◆ **難易度 (difficulty) パラメタ**

2 パラメタ・ロジスティックモデルや 1 パラメタ・ロジスティックモデルを用いる場合、各試験問題の難易度パラメタは、受験者の正答確率が 50% になるところの能力値を意味する。

◆ **識別力 (discrimination) パラメタ**

識別力とは、試験問題が受験者の能力の高低をどの程度敏感に捉えることができるかを示す指標である。通常、個々の試験問題を見ても、テスト全体の得点 (estimated score) が高い受験者の正答率は高く、得点が低い受験者の正答率は低くなるが、受験者の能力値によって正答確率が大きく変化するのであればその試験問題の識別力は高く、あまり変化しないのであれば識別力は低いということになる。一般的に、試験問題には一定程度の高さの識別力が求められる。

- IRT に基づく試験の実施方式には様々なものがあるが、代表的な実施方式としては以下のようなものがある。

◆ **リニア方式**

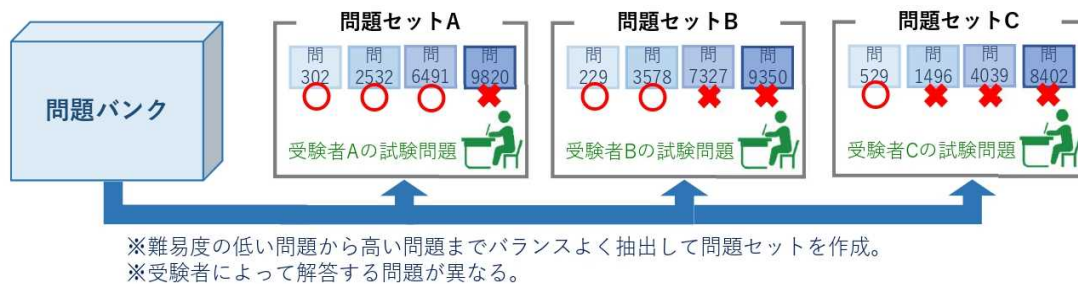
TOEFL iBT® テスト³⁵、TOEIC® Program、医療系大学間共用試験 CBT 等において採用されている方式である。統計的品質のそろった試験問題セットをあらかじめ複数作成し、試験の実施日 (時) や受験する国・地域の異なる受験者集団に対して異なる試験問題セットを使用して実施する。各試験問題セットの品質がそろっているため、試験問題セットの違いによる有利

³⁴ IRT モデルとしては、例えば、以下のようなものがある。

- ・ 1 パラメタ・ロジスティックモデル：1 種類の指標（難易度パラメタ）を用いるモデル。
- ・ 2 パラメタ・ロジスティックモデル：2 種類の指標（難易度パラメタ・識別力パラメタ）を用いるモデル。
- ・ 3 パラメタ・ロジスティックモデル：3 種類の指標（難易度パラメタ・識別力パラメタ・当て推量パラメタ（受験者が偶然に正答できる確率））を用いるモデル。
- ・ 部分採点モデル：試験問題の正誤だけでなく部分点も扱うモデル。

³⁵ TOEFL iBT® テストでは、リーディング・セクションとリスニング・セクションにおいては IRT に基づいて得点の等化がなされているが、スピーキング・セクションとライティング・セクションにおいては、IRT ではなく等パーセンタイル等化法に基づく手法により等化がなされている。

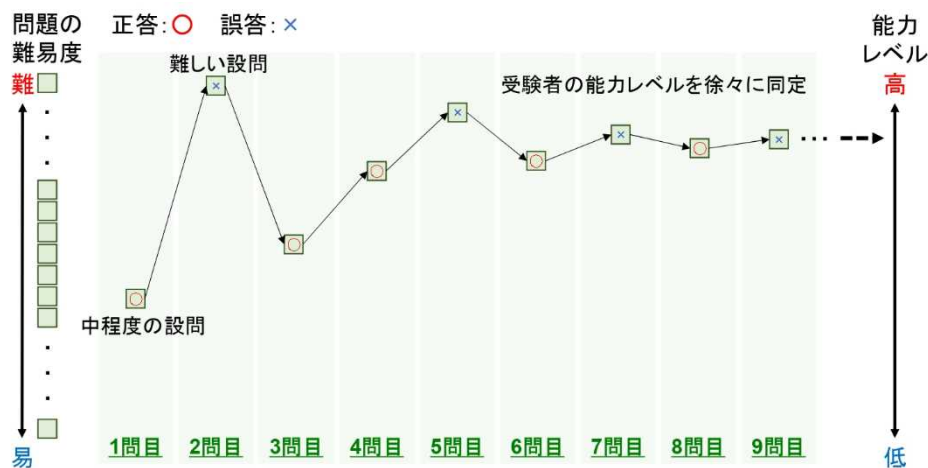
不利は基本的に生じない。



【図6】リニア方式

◆ アダプティブ方式（コンピュータ適応型テスト・Computer-adaptive Testing ; CAT）

GMAT や TOEFL CBT®テスト（1998～2006 年に実施）³⁶において採用されている方式である。具体的には、一人一人の受験者について、1 問ごとの解答の正誤に応じて能力を推定し、その受験者の能力測定の観点から最適な問題（例えば、正答確率が 50%ほどの、解けるか解けないかが不確実であるような問題）を出題する（このため、通常、正答すれば次により難しい試験問題が、誤答すれば次により易い問題が出題される。）。受験者集団の能力差が大きい場合でも、個々の受験者ごとに難易度がカスタマイズされた形で試験問題を出題するので、より少ない出題数で能力を測定できるというメリットがある。



【図7】アダプティブ方式

- IRT に基づいて試験を実施する際は、同一時刻に一齐実施する試験と異なり、本番の試験で出題する前に予備調査を行い、試験問題を本番の受験者に近い仮想受験者に解答してもらい、試験

³⁶ TOEFL CBT®テストの後継である TOEFL iBT®テストでアダプティブ方式を採用しなかった主な理由の一つに、試験後に人間が採点をする問題（新たに加わったスピーキング・セクションなど）が含まれるため、試験中に能力値を推定して適切な難易度パラメタの問題を出題することが困難になったことが挙げられる。

問題の品質をあらかじめ推定した上で、本番の試験を実施することになる。これを行うには、

- ・ 出題される問題が特定されないよう、多数の試験問題を用意する
- ・ 同じ試験問題を本番の試験で何度も利用するという設計の場合、試験問題を非公開とする³⁷といったことなどが必要になる。

- また、IRT に基づく試験では、統計的品質管理を行った試験問題を大量に蓄積したデータベースを構築する場合が多い。このデータベースは「問題バンク」(item bank) と呼ばれる。試験実施時は、この問題バンクから、あらかじめ定められたルールに従って試験問題を選んで出題する。

【コラム⑨】 2パラメタ・ロジスティックモデル

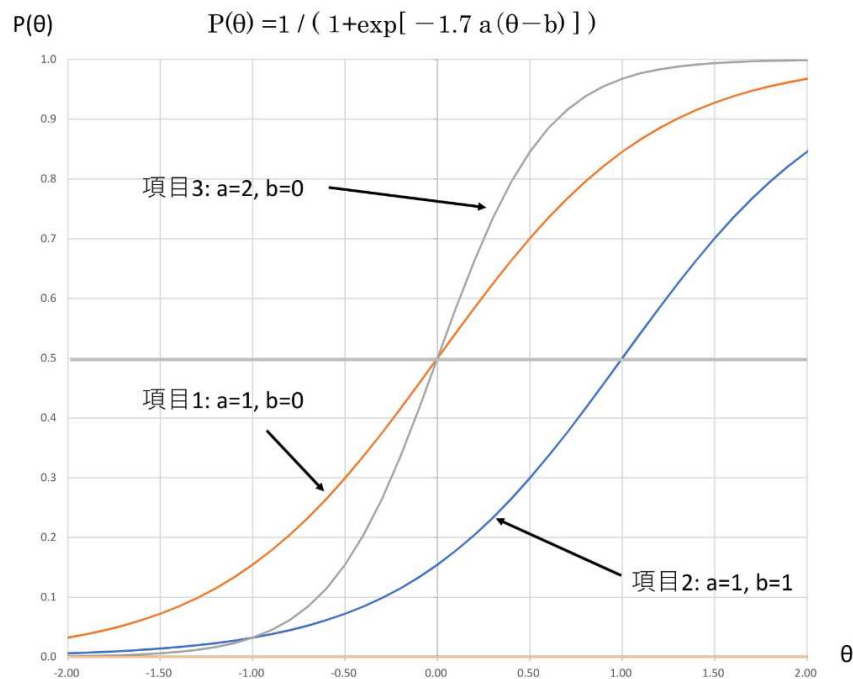
IRT では試験問題への受験者の反応の最小単位（例えば、共通テストであれば、各大問に含まれる小問の解答枠）を項目 (item) と呼び、試験問題は項目の集合であると考ええる。そして、これらの項目は全て、受験者のある一つの特性 (trait；ここではそれを能力と呼び、 θ という記号で記す。) だけを測っていると考える³⁸。

これらの項目の統計的品質は項目パラメタと呼ばれる数値で表現される。項目パラメタの数は用いられるモデルの種類により異なるが、ここでは、項目の識別力を表す a パラメタと項目の難易度を表す b パラメタの二つのパラメタを持つ 2 パラメタ・ロジスティックモデルと呼ばれるものについての説明を行う。

【図コラム⑨－１】に、共に受験者のある一つの特性を測っていると想定される三つの項目の項目特性曲線 (item characteristic curve, 項目反応曲線とも呼ばれる。) を示した。項目特性曲線は、受験者の能力が高くなるにつれ各項目への正答確率が増加していく様子を示したものである。

³⁷ 現行の共通テストの試験問題については、受験者に問題冊子の持ち帰りが認められているほか、試験翌日には多くの新聞に掲載され、後日、大学入試センターのウェブサイトにも掲載されるなど、広く公開されてきた。

³⁸ 各項目が受験者のただ一つの特性のみを測っているという仮定を一次元性の仮定と呼ぶ（【コラム⑬】）。



【図コラム⑨－１】項目特性曲線

【図コラム⑨－１】から明らかなように、能力が高くなればなるほど、どのような項目にも正答する確率が増加していき、最終的には（グラフの右部分で）全ての項目の正答確率は 1 に近づく。また、能力が低くなれば正答確率は減少していく。能力を表す θ は理論的には $-\infty$ から $+\infty$ の値をとるが、図には -2 から 2 までの範囲を示している。

【図コラム⑨－１】で、例えば、項目 1 と項目 2 の項目特性曲線を比べると同じ θ の値でも項目 1 の正答確率の方が高くなっていることが分かるが、このことは項目 1の方が易しい項目であることを示している。これを項目パラメタの違いで見ると、二つの項目の a パラメタの値は共通であるが項目 2 の b パラメタの値が項目 1 の b パラメタの値より大きい。この b パラメタの値は図中に縦軸の 0.5 のところに引いた横線と項目特性曲線との交点の θ の値であり、正答確率が 5 割となる θ の値を意味しているが、その意味で、 b パラメタを「難易度パラメタ」と呼ぶ。

また、項目 1 と項目 3 を比べると、 b パラメタの値は 0 で共通であるが項目 3 の a パラメタの値が大きく、それが、項目特性曲線の θ が 0 付近での傾きの大きさに反映されている。これは、項目 1 では θ が 0 から 1 へ変化しても正答確率の変化は 0.35 程度であるが、項目 3 の場合は 0.45 も変化するというように、 θ の変化に対する正答確率の変化の大きさ（感度）を表す指標であり、 a パラメタは「識別力パラメタ」と呼ばれている。言い換えれば、識別力が高い項目ほど、小さな θ の増加であっても敏感に反応して正答確率が増加する。

なお、項目特性曲線は a パラメタと b パラメタの値と指数関数 $\exp[]$ を用いて

$$P(\theta) = 1 / (1 + \exp[-1.7 a(\theta - b)])$$

というロジスティック関数の形で示すことができる。

項目の識別力が異なる場合、項目の難易度の順番が受験者の能力に依存して変化する場合があります。例えば、項目 1 と項目 3 の難易度パラメタ (b パラメタ) の値は同じであるが、項目特性曲

線を比べると能力の上位層にとっては項目3の方が易しく、能力の下位層にとっては項目1の方が易しい。

(執筆：前川眞一（(独)大学入試センター特任教授，東京工業大学名誉教授))

(2) 異なる試験問題に解答した受験者同士の能力の比較

- 異なる試験問題セットに解答した受験者同士の能力を比較する際に，試験問題セット間の統計的品質の差異に対して何の手立ても講じなければ，不公平が生じる可能性がある。
- この不公平を排除するために，異なる試験問題セットを受験した各受験者のテスト結果を共通の尺度上の得点で表現し，複数の試験問題セットの受験者間で得点を相互に比較することを可能にする統計的処理を行う。この統計的処理のことを「等化 (equating)」という。
- IRT に基づく試験の得点の表示に当たっては，各設問に配点は設定せず，各設問の項目パラメータを用いた IRT の数式により受験者の能力値を推定し，それを基に得点を算出するという方法を採用することが多い³⁹（【コラム⑩】）。

【コラム⑩】 IRT に基づく試験の得点の算出方法

本コラムでは，各項目に配点は設定せず，受験者の各項目（例えば，共通テストであれば，各大問に含まれる小問の解答枠）への反応パターン（正答したら○，誤答したら×）を基にして，項目の難易度パラメータと識別力パラメータを使って受験者の能力値 θ を算出する場合の得点の算出方法を紹介する。能力値の定め方には様々な方法があるが，ここでは，最尤推定法 (maximum likelihood estimation method) と呼ばれる能力値の推定方法を説明する。

例えば，3問の項目から成る試験の正誤反応は，全問正答と全問誤答を除けば⁴⁰，【表コラム⑩－1】に示した S1 から S6 の 6 パターンとなる。例えば，項目1にのみ正答し他の項目には誤答した場合は「S1 ○××」で表される反応パターンを得る。そして，それぞれの反応パターンに対応し θ の値が定まる。

【表コラム⑩－1】 3問の項目から成る試験の正誤反応のパターン

反応パターン	S1	S2	S3	S4	S5	S6
	○××	×○×	××○	○○×	×○○	○×○

³⁹ 難易度等のそろった等質な試験問題セットを使用する場合は，様々な条件を考慮した上で，現行の共通テストのように正答した設問の得点を足し上げた点数（素点）により示すことも可能である。

⁴⁰ 全問正答・全問誤答時の能力値は理論上無限 ($-\infty$ や $+\infty$) となり推定できないため，本コラムの説明には含めない。なお，実際の応用場面では，全問正答・全問誤答の受験者に与える能力値をあらかじめ定めておくケースもある。

項目特性曲線を用いれば、能力が θ の受験者がこれらの反応パターンを得る確率を計算することができる。すなわち、各項目への正誤反応は独立であることを仮定すれば、例えば、S1のような反応パターンをする確率は、能力値が θ の受験者の項目1に正答する確率と、項目2に誤答する確率、そして項目3に誤答する確率を掛け合わせたものとして得られる。いま、各項目への項目特性曲線の値を $P1(\theta)$ 、 $P2(\theta)$ 、 $P3(\theta)$ とすれば、この確率は $P1(\theta) \times (1-P2(\theta)) \times (1-P3(\theta))$ となる⁴¹。

この確率は、それを θ の関数と見た場合に尤度 (likelihood) と呼ばれるが、それを θ の値が-0.5, 0, 0.5の場合に計算したものが【表コラム⑩-2】である。

【表コラム⑩-2】特性値の尤度と最尤推定値

反応パターン	尤度			最尤推定値
	$\theta = -0.50$	$\theta = 0.00$	$\theta = 0.50$	
S1 ○××	0.23	0.21	0.08	-0.3
S2 ×○×	0.04	0.04	0.01	-0.3
S3 ××○	0.10	0.21	0.18	0.1
S4 ○○×	0.02	0.04	0.03	0.1
S5 ×○○	0.01	0.04	0.08	0.7
S6 ○×○	0.04	0.21	0.41	0.7

これらの θ に対応する項目特性曲線の値 (正答確率) とそれを1から引いた ($1-P(\theta)$) の値 (不正解の確率) は、【図コラム⑨-1】から読み取り、【表コラム⑩-3】に示した。例えば、「S1 ○××」という反応パターンを得る確率は、 $\theta = -0.50$ の時には $0.30 \times 0.93 \times 0.85 = 0.23$ であり、 $\theta = 0$ の場合には $0.50 \times 0.85 \times 0.50 = 0.21$ となる。

【表コラム⑩-3】 θ に対応する項目特性曲線の値 (正答確率) と ($1-P(\theta)$) の値 (不正解の確率)

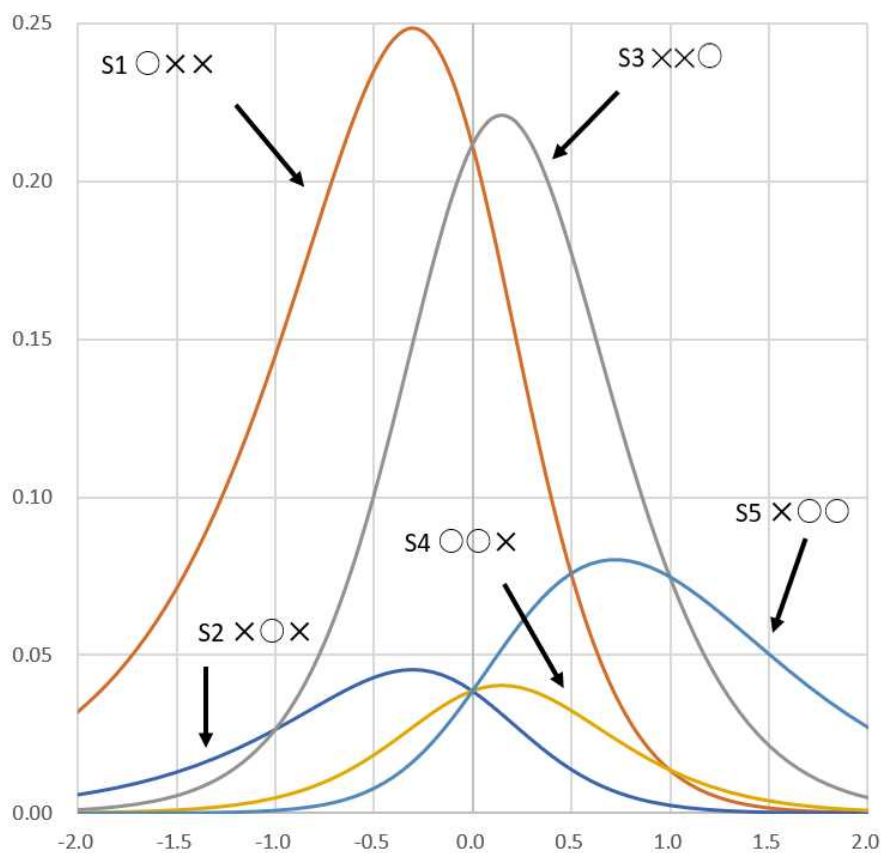
		$\theta = -0.50$	$\theta = 0.00$	$\theta = 0.50$
P(θ)	問題1	0.30	0.50	0.70
	問題2	0.07	0.15	0.30
	問題3	0.15	0.50	0.85
1-P(θ)	問題1	0.70	0.50	0.30
	問題2	0.93	0.85	0.70
	問題3	0.85	0.50	0.15

これを $-2 < \theta < 2$ の範囲で全て計算しグラフにしたものが【図コラム⑩-1】の「S1 ○××」のグラフであり、この反応パターンに対応する θ の尤度関数 (likelihood function) と呼ばれる。

⁴¹ この、各項目への正誤反応が独立であるという仮定を局所独立性の仮定と呼ぶ (【コラム⑬】)。

る。この尤度関数を見ると、 $\theta = -0.3$ 付近でその値が最大となるが、この、最大の正答確率を与える θ の値を持って、反応パターン「S1 ○××」に対応する能力値の推定値とする方法を θ の最尤推定値 (maximum likelihood estimate) と呼ぶ。

もう一つの例として「S5 ×○○」を考えてみると、能力値が θ の受験者がこの反応パターンを得る確率は、項目 1 に誤答する確率 × 項目 2 に正答する確率 × 項目 3 に正答する確率であるから、例えば、 $\theta = -0.50$ の場合には $0.70 \times 0.07 \times 0.15 = 0.01$ であり、 $\theta = 0$ の場合には $0.50 \times 0.15 \times 0.50 = 0.04$ である。そして、それを全ての能力値で計算したものが【図コラム⑩－1】の「S5 ×○○」のグラフであり、その最大値 (最尤推定値) はおよそ $\theta = 0.7$ 付近である。



【図コラム⑩－1】尤度関数

【表コラム⑩－2】には、 θ のおよその最尤推定値を示してあるが、この表から以下のことが分かる。まず、最尤推定値は単純な項目正答数とは一対一に対応していない。例えば、S1, S2, S3 はそれぞれ 1 問だけ正答しているパターンであるが、S3 に対応する θ の推定値の値が大きい。また、S4, S5, S6 に関しても、同じ 2 問だけ正答しているにも関わらず S4 に対応する θ の推定値が低い。また、1 問だけ正答している S3 と 2 問正答している S4 に対応する θ の推定値が同じ値となっている。これは、IRT の 2 パラメタ・ロジスティックモデルにおける θ の推定値が、単純な項目正答数の関数ではなく、識別力パラメタ (a パラメタ) で重みをつけた重み付き正答数の関数として計算できることに由来している。上記の例でいえば、項目 3 の a パラメ

タの値は 2 であるため、S1 から S6 の重み付き正答数は 1, 1, 2, 2, 3, 3 となるが、識別力の高い項目 3 に正答することは、項目 1 と 2 の二つに正答することと同等として取り扱われていることになる。

次に、S1 と S2 を比べてみると難しい項目である項目 2 に正答した S2 に対応する θ の値が大きめに推定されても良いような気がするかもしれないが、この二つの反応パターンに対応する重み付き正答数は同じであるため、 θ の値は等しく推定されている。この点、S2 という反応パターンは、易しいはずの項目 1 にも正答できなかった受験者の反応と解釈すれば納得がいくかもしれない。なお、S1 は易しい項目に正答、難しい項目は誤答という反応パターンであり、逆に S2 は難しい項目に正答、易しい項目には誤答という反応パターンであるため、S2 は起こりにくい反応パターンであると考えられるが、そのことは【表コラム⑩-2】に示した S2 の反応確率の低さに反映されている。

また、最尤推定法で与えられるのは能力値 θ の点推定値であるが、必要とあれば、推定の誤差 (error) を考慮した、一定の幅をもって推定される区間推定値を与えることもできる。たとえば、S3 という反応パターンに対応する θ の点推定値は 0.1 であるが区間推定としては $(-0.1 \leq \theta \leq 0.3)$ という様に表現することも可能である。その際の区間の幅は、試験を構成する項目数が多いほど、また各項目の識別力が平均的に高いほど狭くなることが知られている。

なお、たとえば最尤推定法を用いて反応パターンに対応する θ が得られた場合、それをそのままの形で成績として用いると、例えば、 $\theta=0$ という結果はあたかも「能力がない」と誤った解釈を受験者にされたり、また、 θ が負であるとやはり解釈上誤解や違和感を受験者に与えたりするおそれがある。そこで、例えば、線形変換により、能力値 $\times 100 + 500$ と変換すれば、得点は以下の【表コラム⑩-4】のように示される。

【表コラム⑩-4】能力値と得点

能力値 θ	-4	-3	-2	-1	0	1	2	3	4
得点	100	200	300	400	500	600	600	800	900

(執筆：前川眞一 ((独)大学入試センター特任教授，東京工業大学名誉教授))

2. IRT に基づく共通テストの実施方法と主な課題

- 仮に共通テストを IRT に基づいて実施すれば、1. で述べたような IRT の特徴により共通テストを以下のような形で実施することが可能となる (第 2 章 2. ③で詳述)。
 - ・試験日時の複数回設定が可能
 - ・一人の受験者による複数回受験
 - ・受験者の能力の経時的な変化の把握
- 一方、IRT に基づいて共通テストを実施することを目指す場合、以下に挙げるような論点につ

いて検討が必要である。

(1) 問題作成

- 第1章1.(3)にもあるとおり、現行の共通テストでは6教科 30 科目の試験問題を本試験用、追試験用の2セット作成している。科目ごとに設置されている問題作成部会(約 460 人の大学等の教員)が、年間 50 日程度、大学入試センターのセキュリティエリア内で問題作成の会議を行っている。また、問題点検第一部会(約 160 人の大学等の教員及び高等学校関係者)が、年間 15 日程度、大学入試センターのセキュリティエリア内で試験問題の点検を行っている。
- IRT に基づいて試験を実施する場合には、基本的に、試験問題を事前の予備調査で、本番の受験者に近い仮想受験者に解答してもらうことが必要になるため、本番の試験で出題される問題が特定されないよう、1 科目当たり通常数千～数万問の試験問題を用意する必要がある。作成する問題が大幅に増えるため、現行の共通テストの試験問題と同等の質の維持が求められれば、大量の問題作成を支える十分な数の問題作成者の確保が必要となり、確保できない場合、試験問題の質が落ちる又は問題作成が不可能に陥ることになる。また、問題作成者が多くなればなるほど、問題作成に関する細目(item specification)を緻密に作成し、試験問題に求められる条件を確実に事前に共有することで、誰が問題作成を担当しても、ある程度均質な問題を作成できる体制作りが重要になってくる。
- その場合、大学等の教員を中心とした現行の問題作成体制とは異なる、大学等の教員以外の人材を参画させる新たな体制を構築することも考えられる。例えば、①アイテム・ライター(問題素案作成を作成し、それを試験問題として適切な形に整える者)を専属で大学入試センターで雇用すること、②問題作成業務を個別に委託すること、などが考えられる。しかしながら、こうした従来と大きく異なる方法を導入することについては、大学が共同して実施する共通テストの在り方として、国民的な議論が必要である。
- また、予備調査の実施には、本番の受験者に近い仮想受験者を相当数集める必要があるが、広い学力層が受験する共通テストの場合、それに応じた幅広い能力分布をもつ仮想受験者を用意する必要がある。そのような集団を集めることの実現可能性や、予備調査実施時に生じ得る試験問題の漏洩の可能性等についても考慮する必要がある。
- 大量の問題を作成して予備調査を行うには、多額の経費も必要となる。CBT 問題作成ワーキ

ンググループ⁴²（以下「問題作成 WG」という。）の報告⁴³によると、「情報Ⅰ」を CBT で実施するための試験問題を 5 年間で 1 万問作成する場合、1 年当たり 1 億 125 万円、5 年で総額 5 億 625 万円を要すると試算されている（【コラム⑪】）。

- 大学入試センターの令和 2 年度予算において、PBT で実施する共通テスト 1 科目の問題作成に要する経費は平均すると約 4,100 万円⁴⁴である。したがって、単純計算すると、5 億円をかけて作成した問題を約 12 年間使用し続けることができれば、合計のコストでは現状とほぼ変わらない。しかしながら、IRT の場合には、試験を何回か実施する中で、試験問題の曝露（exposure；試験問題が受験者の目に触れること）や漏洩により項目パラメタが影響を受けてしまうため、試験問題の再利用には限界がある。試験の測定精度を維持するためには、項目パラメタが変動した試験問題を廃棄し、新しい試験問題を問題バンクに追加していくことが必要となることから、実際には、上記問題作成 WG の報告以上の経費が必要になってくると考えられる。
- 特に、共通テストはハイステークスな使われ方をしている試験であることから、問題漏洩のリスクも高く、一方では、試験の精度を保つ必要性も高い。このため、新たに試験問題を作成して、実際に使用する試験問題を入れ替えていく必要性は、他の試験以上に高いと考えられる。加えて、共通テストは高等学校学習指導要領を踏まえて実施することとされているが、学習指導要領はおおむね 10 年に一度改訂されており、改訂時にはその改訂内容を踏まえた対応も必要となるため、教科・科目の再編状況によっては、多くの問題を入れ替えることが求められる可能性がある。

<主な課題>

- 作成すべき問題数が大幅に増える。試験の実施方法や受験者数によっては 1 科目数千～数万問の問題を作成することが求められる。
- 同じ試験問題を本番の試験で何度も利用するという設計の場合、試験問題の曝露や漏洩への対応で頻繁な問題入替え・追加が必要。

<必要な対応>

- 十分な数の問題作成者の確保、又は大学等の教員以外の人材を参画させる新たな問題作成の体制の構築が必要。
- 実施経費の増加に伴う財政負担について検討することが必要。

⁴² CBT 活用検討部会の下に設置された有識者会議。共通テストにおいて教科「情報」の試験を CBT で実施する場合、特に問題バンクを構築して CBT-IRT で実施すると仮定した上で、試行的に問題バンクを構築し、主に問題作成の生産性を中心にそのフィージビリティ（実現可能性）を検証した。

⁴³ 付録 2 「問題バンク構築に係る調査研究について～CBT-IRT での共通テスト「情報」の問題作成に係るフィージビリティの検証～（報告）」。

⁴⁴ 大学入試センターの令和 2 年度予算において、PBT で実施する共通テスト（6 教科 30 科目）の試験問題作成に要する経費として約 12.3 億円が計上されており、単純計算で 1 科目当たりの平均試験問題作成経費は 4,100 万円となる。

【コラム⑪】問題バンクの構築・メンテナンス

本文(1)では問題バンクについて触れたが、そもそも問題バンクはどのように構築するのか、また、一度構築した後に、どのようなメンテナンスが必要なのだろうか。

○必要問題数の算出

問題バンクには、一般に数千～数万問の試験問題を蓄積することが求められるとされるが、実際には、各試験において蓄積すべき問題数は、試験をめぐる状況に応じて異なり、ケース・バイ・ケースでの算出が必要である。

問題バンクに蓄積すべき問題数を変動させる要因としては、①一つ当たりの試験問題セットに含める問題数、②作成すべき試験問題セットの数、③出題領域の細分化の度合いなどが挙げられる。また、後述するように、試験問題（案）に対する予備調査を実施する場合、④問題の採択率（予備調査で使用した問題のうち基準を満たしたものの割合）や、⑤問題を何回繰り返して使用するか（どの程度なら繰り返し使用できるか）等についても考慮する必要がある。

○試験問題（案）の作成

問題作成体制を整えた上で大量の試験問題（案）を作成する。なお、次に述べるように、予備調査のデータを基に推定した難易度や識別力の値が統計的な基準を満たしたもののみを問題バンクに登録することになるため、実際には、問題バンクへの登録が必要な試験問題数を超える試験問題（案）を作成する必要がある。

なお、出題領域や下位分類を設定する場合には、試験問題（案）の作成時にメタ情報（出題領域や下位分類を示すラベル）を付与することで、試験問題の効率的な管理が可能になる。

○試験問題（案）の項目パラメタの推定

次に、作成した試験問題（案）の項目パラメタを推定するための予備調査が必要になる。このための方法としては、以下のように、本番の受験者に近い仮想受験者を対象にした「プレテスト⁴⁵⁾」の実施や、本番の試験において新作の試験問題（案）をダミーとして出題するなどがある。

◆ プレテストの実施

プレテストは、本番の受験者に近い仮想受験者に協力を求め、試験問題の項目パラメタをあらかじめ推定する方法である。多数の問題について一定数以上の仮想受験者の解答を依頼することになるため、複数種類の試験問題セットを用いて、効果的な解答データ収集を計画することが必要である。

このプロセスは、いわば、仮想受験者が手分けをして解答を行うイメージとなるが、多数の問題に対してプレテストを行うため、同一の試験問題セットに含まれる問題の組み合わせ方には留意が必要である。例えば、互いにヒントとなりそうな問題ペアがあった場合に

⁴⁵⁾ フィールドテスト、試行テストとも呼ばれる。

は、別々の問題セットに入れることが望ましい。反対に、同時に出题すべき問題ペアについては、同一の試験問題セットに含めることが望ましい。なお、この方法を用いる場合、実際の試験が行われる前の段階で、相当数⁴⁶の仮想受験者が問題の内容を知ることになる。もちろん、誓約書等により機密保持を約束させる、書類等の持ち出し等を認めないといった措置を講じることは考えられるが、「記憶による持ち出し」を完全に防ぐことはできないため、試験問題に関する情報が一定程度漏洩する可能性は否定できない。

◆ 本番の試験で新作試験問題をダミー試験問題として出题

本番で解答する受験者の得点算出用の試験問題に、「ダミー試験問題」として新作試験問題を混ぜて解答させ、その解答データから難易度を推定するという方法もある。この方法の利点は、仮想受験者ではなく本物の受験者から、試験問題解答へのモチベーションが最も高い状態で、新作試験問題に対する解答データを得られる点である。一方で、従来の日本の試験の多くでは、与えられた試験時間をどのように使うかについても、測定する能力の一環として捉えられてきた部分があることから、仮に、ダミー問題を加える場合には、試験時間の在り方について慎重な検討を要する。

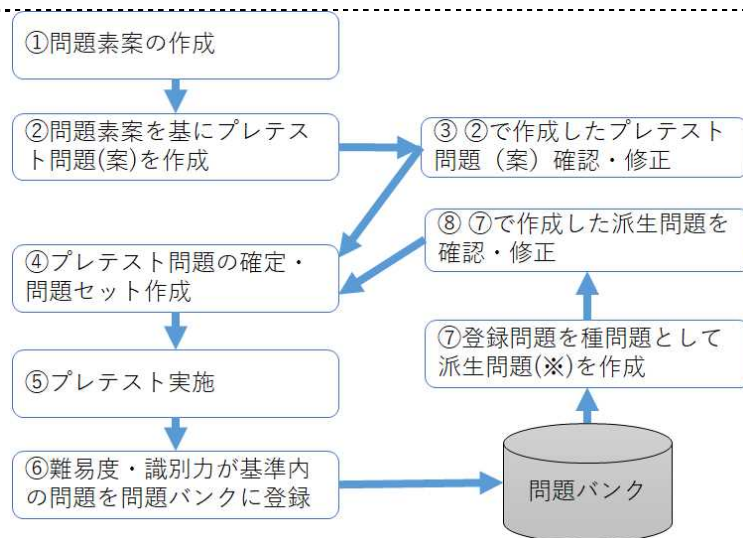
各試験問題（案）の解答データの収集・分析後、項目パラメタの値が条件を満たすもののみ、項目パラメタの値を付与した上で問題バンクに蓄積する。項目パラメタの値が条件に合わない（難易度パラメタが高すぎる又は低すぎる、識別力パラメタが低すぎる等）場合には、該当する試験問題（案）を廃案にする、再修正を施して改めて予備調査を行うなどする。

○問題バンクのメンテナンス

問題バンクは一度構築した後も継続的なメンテナンスが必要である。これは、試験を何回か実施する中で、試験問題の曝露や漏洩によって項目パラメタが変動するためである。試験の測定精度を維持するためには、項目パラメタが変動したと考えられる試験問題を廃棄し、新しい試験問題を問題バンクに追加していくことが必要となる。また、学習指導要領の改訂など、試験で測定しようとする能力や問う内容が変わった場合にも、新しい試験問題と入れ替える必要がある。

問題作成 WG においては、共通テスト教科「情報」の問題バンクを構築する場合の問題作成プロセスや問題作成体制を具体的に想定し、そのプロセスや体制を、小規模ではあるが実際に取り組んでみたところである。

⁴⁶ どの IRT モデルを用いるかに応じて異なるが、少なくとも 1 問当たり数百人分の解答が必要となる。



【図コラム⑪－１】問題作成WGで取り組んだ問題バンク構築の流れ

この取組と同様の方法で５年間にわたって１万問の試験問題を問題バンクに蓄積する場合、問題作成や予備調査（ここではプレテストを想定）の実施などのため、【表コラム⑪－１】にあるような人員・経費が必要となる。必要な経費については、１年当たり１億１２５万円、５年間総額で５億６２５万円と試算されている。

【表コラム⑪－１】１万問の試験問題を問題バンクに構築する場合に必要な人員・経費

	必要人数〔年〕	必要経費〔円／年〕
問題素案作成（①）	50（各作成者が作成する素案数：20問） ⁴⁷	0.5万×20問×50人＝500万
第一部会（②）	20人（会議日数：50日）	50万×50日＝2,500万
第二部会（③）	20人（会議日数：50日）	50万×50日＝2,500万
問題管理委員会（④）	10人（会議日数：25日）	25万×25日＝625万
予備調査（プレテスト）（⑤）	参加者 20,000人 ⁴⁸	0.2万×20,000名＝4,000万

※表中の丸数字は、対応する【図コラム⑪－１】のプロセスを指す。

【コラム⑫】医療系大学間共用試験の問題作成について

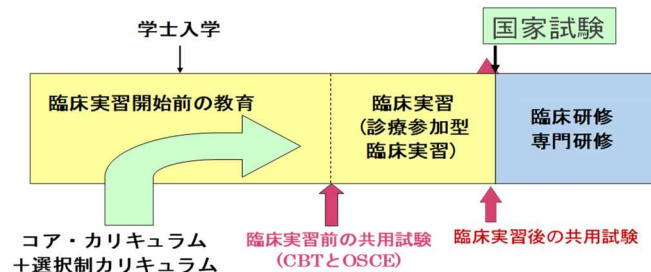
医療系学部（医学部、歯学部）教育については、その到達目標を示した「モデル・コア・カリキュラム」が国において策定されており、各学部ではこれに基づいてカリキュラムが編成されて

⁴⁷ 各作成者が作成できる素案数は、科目の特性や出題形式により異なる。

⁴⁸ プレテスト参加者数は、

- ・ 1問当たり延べ200人の参加者が必要
 - ・ 参加者1人が一つの試験問題セットに含まれる25問を解答
- という仮定に基づいて算出した。

いる。医療系大学間共用試験（以下「共用試験」という。）は、臨床実習に参加する学生の知識と技能・態度がモデル・コア・カリキュラムに定める水準に到達していることを確認するための大学間共通の評価システムで、CBT-IRT で実施されている。



【図コラム⑫-1】医師養成における共用試験の位置付け

大学間共通で臨床実習開始前の学生の到達度を確認する強いニーズがあった一方、学修進度や臨床実習開始時期が大学により異なり、到達度を確認する試験を全ての大学で同時期に行うことはできなかったため、異なる時期・場所でも公平に実施・評価できる試験が必要とされた。このため、共用試験のうち知識に関する部分の試験はCBT-IRTで実施することとなったのである。

ハードウェアの使用方法はLAN方式⁴⁹で、試験実施機関である医療系大学間共用試験実施評価機構（CATO）⁵⁰のサーバに試験問題が配信され、各大学のパソコンやネットワーク（LAN）を使用して実施する。試験で使用するハードウェアは各大学が用意し、各大学は実施環境が適切かを動作確認キットで確認する。受験者数は医学系・歯学系合わせて年間約11,000人、受験料は令和2年度時点で25,000円である。

共用試験の出題は、CATOで構築する問題バンクから行われる。以下で共用試験の問題バンクの構築やメンテナンスについて紹介する。

○試験問題の概要

受験者は以下の6ブロック（各ブロック60分）、計320設問の試験問題を解答するが、320問のうち約80問がダミー試験問題（解答データを得ることを目的に本番の試験問題に混ぜる新作試験問題で、解答内容は試験の成績には反映されない）⁵¹である。成績表示は「合格」「不合格」の2種類である。

ブロック1～4：単一多肢選択形式60設問

ブロック5：多選択肢連問形式40設問（鑑別診断、病態等）

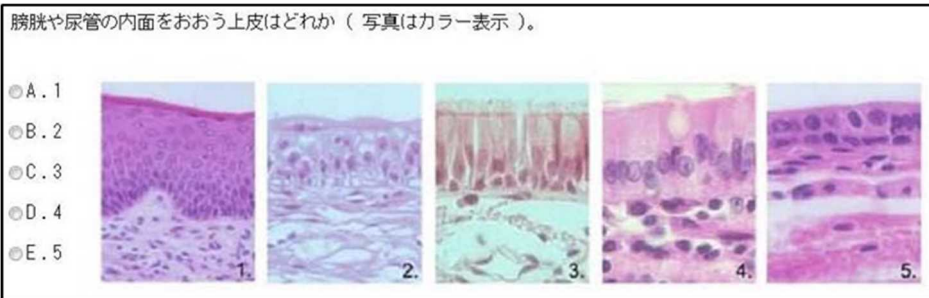
ブロック6：順次解答4連問形式40設問（臨床推論）

⁴⁹ LAN方式については、第3章1. (2)で詳述。

⁵⁰ 試験実施業務を支える組織体制の確立や、試験実施用システムの開発、試験問題の蓄積、学生の成績と試験問題のセキュリティ確保、試験監督の派遣等のために、試験実施機関として医療系大学間共用試験実施評価機構（CATO）が設置された。

⁵¹ ダミー試験問題については、【コラム⑪】で詳述。

例題 8 B-1-(2)-①-1) 上皮組織と腺の構造と機能を説明できる。



例題 47 E-2-(2)-6) 生化学的検査項目を列挙し目的、適応と異常所見を説明し、結果を解釈できる。

食事の影響を受けやすい血液生化学検査はどれか。

- A. 尿酸
- B. 総蛋白
- C. ナトリウム
- D. クレアチニン
- E. トリグリセリド

例題 53 E-3-(5)-③-1) 胸部診察で確認すべき項目を列挙し、視診、触診、打診と聴診ができる。

心臓の聴診上、I 音と同時に開始し II 音まで続く雑音を聴取した。この心雑音はどれか。

- A. 連続性雑音
- B. 収縮中期雑音
- C. 全収縮期雑音
- D. 収縮後期雑音
- E. 拡張早期雑音

【図コラム⑫－２】共用試験（医学系）の公開問題の例

※共用試験のその他のサンプル問題は**参考資料 4**を参照

○試験問題（案）の作成

問題バンクを作成するために、平成 14～17 年の 4 年間にわたりトライアルを行った。問題作成依頼は CATO から問題入力システムなどを含む問題作成キットを提供して参加大学の教員に問題作成を依頼した。4 年間で、医学系 30,000 問以上（参加 80 大学）、歯学系 25,000 問以上（参加 28 大学）の試験問題が提出された。

大量の問題を集めることができた理由としては、受験者層が臨床実習開始前の医療系学部生に絞られており問題を作成しやすかったこと、教員に共用試験の必要性が理解されていたこと、国家試験の作問経験者が多かったことなどが挙げられる。なお、トライアル期間中は問題作成者に対して謝金等は支払われていない。

○問題バンクの蓄積

トライアル期間中医学系約 28,000 人以上、歯学系約 10,000 人以上の学生が、共用試験システムを用いて試験を実施した。試験結果については CATO の事後評価解析委員会により評価が行われ、4 年間で医学系が合計約 12,000 問、歯学系が合計約 8,000 問を問題バンクに蓄積した。

○問題バンクのメンテナンス

難易度に著しい変化が生じた問題、受験者全員がほぼ正解する問題、現在の状況と合わなくな

った問題などを問題バンクから削除する作業を毎年実施している。医学系については、プール問題の管理などを実施する専門部会が1回3人で年80回程度の実施で、問題バンクに蓄積されている問題の評価、新しく問題バンクに入れる問題の評価、事後評価等を行っている。また、不足となったモデル・コア・カリキュラムの領域の問題を新たに募集するなどして、新陳代謝を行っている。

○共用試験の CBT-IRT について

上述のように大変な労力をかけて行われている共用試験が普及してきた一つの理由を示す。平成26年から全国医学部長病院長会議が、共用試験の統一的な合格基準を設定し、合格者に Student Doctor の証明書を発行している。項目反応理論より求めた IRT 標準スコアの数年にわたる安定性、理論的な妥当性などから、この統一的な合格基準（推奨最低合格ライン）に、IRT 標準スコアが採用された。他の試験との相関が高いこと（各大学における報告等）、合格基準周辺の標準誤差が十分に小さいこと（共用試験実施評価機構実施の講演会）が報告されている。また、内容妥当性については、問題作成がモデル・コア・カリキュラムに準拠していることがあり、さらに多くの教員が繰り返し問題のチェックを行っていることから適切な問題が提供されている。このように、十分な質保証を得るためのツールとして IRT 標準スコアが評価されてきている。

（執筆：仁田善雄（公益財団法人医療系大学間共用試験実施評価機構理事））

(2) 試験問題の非公開

○ IRT に基づく試験の得点の信頼性を維持するためには、各試験問題の項目パラメタの変動がないようにする必要がある。しかしながら、試験問題を原則非公開とし、同じ試験問題を本番の試験で何度も利用するという試験設計の場合、試験問題の曝露や漏洩による試験問題の品質の変動が生じ得る。**【コラム⑪】**のとおり、受験者等が記憶に基づいて問題を持ち出すこと、いわゆる「記憶による持ち出し」は防ぎようがなく、意図的あるいは意図しない形での問題漏洩の可能性は否定できない。既存の IRT に基づく他の試験の事例を踏まえても、試験問題を非公開としていても、「記憶による持ち出し」は発生しており、漏洩を完全に防ぐことはできていないというのが実態である。

○ 試験問題の漏洩に対処するための実効的な方策としては、

- ・問題バンクに蓄積する問題数を多くする
- ・問題バンク内の試験問題の品質を常に確認し、曝露や漏洩により品質が変動した疑いのある問題を問題バンクから速やかに引き下げ、新規問題を追加する

などが考えられ、これらを徹底することが求められる。しかしながら、(1)で述べたように、問題の追加や入替は容易ではない上、それらを徹底しても、漏洩による影響を確実にゼロに抑えるこ

とは難しい。

- また、現行の共通テストの試験問題が試験終了後に公開されていることを踏まえると、同じ試験問題を本番の試験で何度も利用するという試験設計とする場合、一部のサンプル問題は公開するとしても、共通テストの試験問題が原則非公開になることについて、受験者や保護者を含む社会全体の理解を得る必要がある。また、共通テストにおいては、「高等学校教育へのメッセージ性」を重視した問題作成を行っているが、試験問題が非公開になることで、試験問題を教育現場で活用することが困難になることについても留意が必要である。
- なお、現行の共通テストの結果も、大学によって様々な形で使われている。共通テストの得点で合否が決まるようなハイステークスな使われ方がある一方で、共通テストの得点だけで合否が決まるのではなく、他の資料等も総合した上で判断するようなローステークスな位置付けである場合には、仮に、受験者が試験問題の一部を事前に知っていても、難易度や識別力等の項目パラメタや解答時間の経時的な変化が極めて小さくなる可能性がある。こうした場合であって、試験の実施や結果の利用に支障がないことが確認できれば、試験問題の漏洩による影響を考慮する必要性が低い場合も考えられる。また、そうした試験の場合には、全試験問題をあらかじめ公開することも選択肢としては考えられる。

<主な課題>

- 同じ試験問題を本番の試験で何度も利用するという設計とする場合、試験問題が原則非公開になる。

<必要な対応>

- 得点の信頼性を維持するためには、試験問題の漏洩を防ぐことが必要（ただし、漏洩の影響をゼロにすることは不可能であり、そのことに対する理解も必要）。
- （一部のサンプル問題を除いては）試験問題を教育現場で活用できなくなることへの理解が必要。

(3) 出題形式（大問か小問か）

- 現行の共通テストは、従来のセンター試験の出題形式を引き継ぎ、原則として大問形式で出題されている。ここでの大問とは、「一つのストーリーに沿って複数の小問⁵²を組み合わせたままり」のことを言う。大問形式による出題は、文脈を把握しながら思考する活動に取り組ませることができるなど、教育学的見地からも価値があると指摘されている。一方、大問形式のまま IRT に基づいて共通テストを実施する場合、次のようなことを検討する必要がある。

⁵² 小問とは、受験者が解答を行う最小の単位である。

- まず、大問形式によっても、局所独立性 (local independence) の仮定 (受験者の能力が同じ場合であれば、一方の試験問題の正誤は他方の試験問題の正誤に影響を及ぼさないという仮定) を満たすことができるかどうかである⁵³。ある試験を IRT に基づいて実施する場合、当該試験の問題に対する受験者の解答データが、局所独立性の仮定を満たしていることが求められるため、ある問題が次の問題を正解するための前提になるなど、問題間での強い連動性がないことが必要になってくる (**【コラム⑬】**)。
- もっとも、大問形式であるからといって、直ちに局所独立性の仮定が失われるわけではなく、受験者のデータを分析した結果も踏まえた総合的な判断が必要である。大問形式を用いたとしても、局所独立性の仮定が十分満たされている場合は、大問内に含まれる各小問について正誤の採点を行い、2 パラメタ・ロジスティックモデル等の 2 値型 (正答と誤答の二つのみが存在する) 項目反応モデルを適用して、能力値を推定することが考えられる (**【コラム⑨】**)。仮に、局所独立性の仮定が満たされないと判断された場合は、大問内の正答数得点 (その大問で何問に正答したか; 大問内の部分点のような扱いになる) を用いて、部分採点モデルや段階反応モデル等の多値型 (採点結果が二通りより多くなる) 項目反応モデルを適用すること⁵⁴なども考えられる。大問形式に対してどのような項目反応モデルを適用するかは、受験者の解答データの統計学的特徴、試験の目的や得点の利用方法等に応じて異なり得るだけでなく、分析者によっても判断が分かれ得るところである。
- また、問題バンクを用いる場合には問題を大量に作成する必要があるため、大問形式の問題は小問形式に比べて大量作成が困難であると予想されるため、どのように問題作成を行うかについて検討が必要である。
- さらに、大問形式の試験問題は小問形式の試験問題に比べて印象に残りやすく、「記憶による

⁵³ 本文では主に局所独立性の仮定に関して説明したが、局所独立性の仮定と次に述べる一次元性の仮定は密接に関連している。大問形式のまま IRT を導入する場合、一次元性 (unidimensionality) の仮定を満たすかについても、より丁寧な検討が必要となる。一次元性の仮定とは、一つの教科・科目の試験に含まれる問題の正誤データの背後に一つの能力次元が想定できるとする仮定のことである。一次元性の仮定は、本来 IRT だけに限定されたものではない (例えば、現行の共通テストにおいても各設問の配点を足し上げて得点を求めることが多く行われるが、一本の数直線上に布置される合計得点を使用するということは、暗黙のうちに設問間の一次元性を仮定していると考えられる)。しかし、大問形式のまま IRT を導入した場合、試験全体の出来不出来を左右する能力次元に加え、各大問で扱われるテーマや測定内容・領域などに起因して大問内の項目の出来不出来を左右する個別の次元が見出される可能性があるため、一次元性の仮定を満たすのかについてより丁寧な検討が求められるのである。

⁵⁴ 例えば、ある大問が四つの小問 (配点各 1 点) から構成されている場合、この大問の得点は、全問誤答の 0 点から全問正答の 4 点までの 5 段階の採点結果をもつことになる。このような大問を 2 値項目と捉えてしまうと、全問誤答と全問正答の二通りの採点結果しか想定しないことになり、適切でない。

持ち出し」が、より容易であると考えられる。しかし、(2)にあるとおり、同じ試験問題を本番の試験で何度も利用するという設計とする場合、試験問題の非公開を前提に実施する必要があるため、「記憶による持ち出し」が容易な形式での出題は、IRT にはなじまない可能性があることにも留意が必要である。

- 以上のように、共通テストを IRT に基づいて実施する場合、従来の出題の在り方にも影響を与える可能性がある。どのような出題形式とするのが適切かについては、様々な見地から十分に検討することが求められる。

<主な課題>

- 現行の共通テストのように大問形式で出題する場合、以下の課題がある。
 - ・問題の大量作成が困難。
 - ・「記憶による持ち出し」が小問形式に比べて容易。

<必要な対応>

- どのような出題形式とするのが適切かについて、様々な見地から十分に検討することが必要。

【コラム⑬】局所独立性の仮定、一次元性の仮定

本文(3)及び脚注 53 で言及した局所独立性の仮定及び一次元性の仮定について説明する。

○局所独立性の仮定

局所独立性の仮定とは、受験者の能力が同じ場合であれば、一方の試験問題の正誤は他方の試験問題の正誤に影響を及ぼさないという仮定である。局所独立の「局所」とは、「受験者の能力が同じであれば」という意味である。また「独立」とは、解答結果に関連性がないという意味である。つまり、一次元として取り出した能力の高低のみが、試験問題間の正誤の間に結び付き(相関・連関)を生じさせるが、それ以外の要因は試験問題間の正誤に相関や連関を生じさせないことを意味する。IRT を使う場合には、局所独立性を仮定する必要があることが一般的である。

局所独立性は、①前の試験問題の答えを使って後の試験問題に解答するような形式を伴う問題がある場合(項目連鎖(item chaining))、②提示された文章内の特定のテーマなど、必ずしも当該の能力の測定とは直接関係がないとされる素材(問題に含まれている図表、文章など)が複数の試験問題間で共有されている場合(文脈依存(context dependence))、③測定を意図した能力とは別の能力次元が試験問題の出来不出来を反映してしまっている場合、に阻害される。

局所独立性の阻害が想定される場合に採られる対応は二つある。一つは、問題作成の段階でこういった構造となることを避ける対応が挙げられる。項目連鎖が生じないような問題や、ある程度一般的に受験者が知っているであろうテーマを題材とした問題を作成することになる。もう一つは、項目連鎖や文脈依存が生じている複数の試験問題を一つの大問(基本単位)と見なして、

多値型の IRT モデルを適用する対応である。

局所独立性の仮定: 受験者の能力が同じ場合であれば、一方の問題の正誤は他方の問題の正誤に影響を及ぼさないという仮定

受験者Aと受験者Bの能力が等しい場合、局所独立性が満たされていれば、能力が等しいので、ほぼ同じ得点となるような正答・誤答状況になる。

	問1 (易)	問2 (普通)	問3 (普通)	問4 (普通)	問5 (難)
A	○	×	○	○	×
B	○	○	○	×	×

しかし、次のような場合、局所独立性が満たされず、受験者Aと受験者Bの得点が異なるケースが生じ得る。

①項目連鎖:

例えば、問2の解答内容を使って問3以降に解答するという構造の問題において、問2に誤答した受験生は問3以降に正答できないという状況のこと。

②文脈依存:

ある共通のテーマに関連させて複数の問題が出題されているとき、受験生が当該テーマについて詳しいか否かで問題の解きやすさが変わってしまうこと。

大問	問1 (易) 問2 (普通) 問3 (普通) 問4 (普通) 問5 (難)					
	問1 (易)	問2 (普通)	問3 (普通)	問4 (普通)	問5 (難)	
A	○	×	×	×	×	問2で誤答した受験者Aは、同等の能力をもつ受験者よりも、正答できるはずの問題で誤答となってしまった。
B	○	○	○	×	×	

大問	問1 (易) 問2 (普通) 問3 (普通) 問4 (普通) 問5 (難)					
	問1 (易)	問2 (普通)	問3 (普通)	問4 (普通)	問5 (難)	
A	○	○	○	○	○	たまたま南米の音楽を聴くことが趣味だった受験者は、趣味の知識で読解力を補うことができ、同等の読解力をもつ受験者よりも正答できた。
B	○	○	○	×	×	

【図コラム⑬-1】局所独立性の仮定について

○一次元性の仮定

一次元性の仮定とは、一つの教科・科目の試験に含まれる問題の正誤データの背後に一つの能力次元が想定できるとする仮定のことである。すなわち、受験者1人につき測定される能力推定値は1種類に限定することを意味している。大規模試験においてIRTを適用している事例の多くは、この一次元性を仮定している。

IRTモデルの中には多次元性に対応したものもある⁵⁵が、選抜を目的とする試験の場合は、一つの試験で複数次元にまたがると運用が煩雑になること等から、一次元に縮約して得点を表示することが多い。

⁵⁵ 二つ以上の能力について考えるときには、補償型（一方の能力が低くても、他方の能力が補って正答できる構造）か非補償型（どちらも高くないと正答できない構造）か等、適切なIRTのモデルについて綿密な検討が必要となる。また、複数の下位能力が想定されていても、解答データを説明できる（より大枠な意味合いをもつ）単一次元の能力にまとめることが適切な場合もあるし、その逆も考えられる。

一次元性の仮定：試験問題の正誤の背後に1つの能力次元が想定できるという仮定

(例)：世界史(政治史, 経済史)の試験において

※問1～4は政治史の問題、問5～8は経済史の問題。問番号下の()は当該問題の難易度。

	問1 (易)	問2 (普通)	問3 (難)	問4 (難)	問5 (易)	問6 (普通)	問7 (難)	問8 (難)
A	○	○	○	○	○	○	○	○
B	○	○	×	×	○	○	×	×
C	○	×	×	×	○	×	×	×

この場合

この試験で測定している能力は一次元と推定される
(＝一次元性の仮定が満たされている)：

受験者の得点は一本の数直線上に位置づけられる

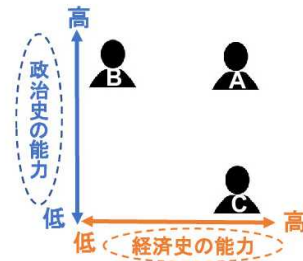


	問1 (易)	問2 (普通)	問3 (難)	問4 (難)	問5 (易)	問6 (普通)	問7 (難)	問8 (難)
A	○	○	○	○	○	○	○	○
B	○	○	○	○	×	×	×	×
C	×	×	×	×	○	○	○	○

この場合

この試験で測定している能力は二次元と推定される
(＝一次元性の仮定が満たされていない)：

受験者の得点は二次元の平面に位置づけられる



【図コラム⑬－2】一次元性の仮定について

(4) 成績の表示方法

- 現行の共通テストでは、正答した設問に割り振られた配点を足し上げた「素点」(raw score)が受験者の得点となり、大学入学者選抜で使用されている。このため、マークの誤り等がない限り、採点結果と自己採点は一致することになっている。
- しかし、IRTに基づく試験の得点の表示に当たっては、一般に、1.(2)で述べたように、各試験問題の項目パラメタを用いたIRTの数式により受験者の能力値を推定し、それを基に得点を算出するという方法を採用することになる(【コラム⑩】)。
- 現行の共通テストや他の日本の試験の多くが素点に基づいて行われていることを踏まえると、共通テストの成績をIRTの数式により算出した得点を用いて表示することとした場合、自己採点結果と得点の関係が分かりにくく、例えば、従来のように受験者が自己採点により自身の成績を予測することが困難であったり、成績開示後に受験者が自分の得点に疑義を抱くという事案が生じたりする可能性がある。このため、受験者や保護者、高等学校関係者等への事前の周知や、成績に疑義が生じた場合の対応方法の検討などを行い、新しい成績の表示方法について受験者や保護者を含む社会全体の理解を十分に得ることが必要である。なお、IRTに基づく場合でも素点

で成績を表示できるような設計とする⁵⁶ことも可能であり、こうした選択肢を視野に入れることも考えられる。

<主な課題>

- IRT の数式に基づいて算出した得点により成績を表示することが多いが、その場合、得点は現行の共通テストで用いられている「素点」とは異なる表示になる。

<必要な対応>

- 「素点」とは異なる表示とすることとした場合、受験者自身が自己採点結果と得点の関係を理解するのが困難になるため、新しい成績の表示方法について受験者や保護者を含む社会全体の理解を十分に得ることが必要。

【コラム⑭】日本的試験文化

近年では大学入学者選抜の在り方も多様化が進んでいるが、それでも、我が国における試験文化の形成に、共通1次と、それに続くセンター試験という大規模共通試験が果たしてきた役割は大きいだろう。その特徴としては、以下の点が指摘できる。

① 年に一度、同一の新作問題を用いて同一時刻に一斉で行われる。

- ・全ての試験場で同一の試験問題が出題される。
- ・試験問題について本番前の予備調査は行わない。

② 試験問題は直後に公開される。

- ・過去に出題された試験問題は教育・受験勉強に利用される。

③ 多肢選択式の大問形式が多く出題される傾向がある。

- ・小問形式より大問形式の方が高度の認知能力を測れるとされる。
- ・各試験問題がどの分野のどのような能力を測っているかが明確ではない。

④ 成績表示には、正答した設問の配点を足し上げた点数である「素点」が利用される。

- ・選抜においては素点が重視される。

このような特徴は、「日本的試験文化」として、大学入試の一つのスタンダードとして社会的な定着を見せている。すなわち、公平・公正な選抜を行うためには、上記の特徴を備えた方式で実施すべきという考え方が浸透しているといえるのではないだろうか。

一方、大学入学者選抜を IRT に基づいて行う場合、例えば、複数の試験問題セット、複数の試験日、試験問題の非公開等、こうした伝統的な日本的試験文化と相容れない部分が多い。

⁵⁶ 1. (2)でも言及したとおり、難易度等のそろった等質な試験問題セットを使用する場合は、様々な条件を考慮した上で、現行の共通テストのように素点により成績を表示することも可能である。

新しい試験を導入する場合には、受験者や保護者を含む社会全体の理解は不可欠である。そのためには、共通テストに IRT を導入する場合は、そのメリット及びデメリットについて丁寧に周知していくことが必要である。

<参考文献>

石井秀宗（2018）「大学入試における共通テストの複数回実施は実現可能かー日本のテスト文化やこれまで見送られてきた理由などからの検討ー」, 名古屋高等教育研究 第 18 号, 23-38.

前川眞一（2014）試験の日本的風土 大学入試センターシンポジウム 2014 大学入試の日本的風土は変えられるか (URL: <https://www.dnc.ac.jp/news/20141203-01.html>)

3. IRT に基づく共通テストにより実現できること

- 2. では、共通テストを IRT に基づいて実施する際にどのような課題があるかについて検討した結果をまとめたが、その中では、実現に向けて克服すべき課題も多いことが明らかになった。本節では、仮にそれらの課題が克服されて、共通テストを IRT に基づいて実施することとなった際に、どのようなことが実現可能になるかを整理する。

(1) 試験の年複数回実施

- IRT に基づいて共通テストを実施すれば、異なる試験問題に解答した者同士の得点も比較できるようになる。そのため、現行では原則として年 1 回のみの試験であるところ、年に複数回実施することが可能となる⁵⁷。これにより、年 1 回の同一時刻一斉実施に付随する、病気や事故等の事情により受験できないといったリスク、今般の新型コロナウイルス感染症などの流行性疾患の感染拡大や大規模な自然災害発生により実施不可能になるといったリスクを一定程度解消できる。
- ただし、共通テストを年に複数回実施することとした場合、どの時期に試験を実施するかが問題となる。現行の共通テストの実施時期より早い時期、すなわち 1 月よりも前に共通テストを実施することについては、高等学校の授業の進め方や学校行事、課外活動等にも大きな影響を与える可能性がある。また、分割実施とする場合、試験日が早いか遅いかによる不公平（感）が生じたり、既に受験し終わった受験者と未受験の受験者が混在する時期に受験者の心理状況等に影響

⁵⁷ また、第 3 章 1. (9) で言及したように、受験環境の整備やトラブルへの対応を考慮すると、共通テストに CBT を導入する場合、同一時刻一斉実施ではなく分割実施（試験日時を複数設定）の方が実施しやすいことにも留意が必要である。

が生じたりすることなどが考えられる。このため、どのような時期・期間で試験を実施するかについては、高等学校教育への影響などを含めて慎重に検討していく必要がある。

(2) 一人の受験者による複数回受験

- 共通テストを IRT に基づいて実施し、試験日を複数設定することとした場合、一人の受験者が複数回受ける「複数回受験」を認める制度設計も可能となる。いわゆる「一発勝負」の試験の場合には、病気や事故等の事情によって受験できなくなる（あるいは、受験はできても、試験問題の内容や当日の体調等に影響される）というリスクがあるところ、複数回受験が可能になることで、受験者にとっても、また大学にとっても、より実力が反映された結果に基づいた選抜を行うことができると考えられる。
- しかし、複数回受験について検討する際には、民間の英語資格・検定試験の活用に関しても議論のあった、経済格差や地域間格差の問題を避けて通ることはできない。仮に、1回受験する度に検定料を徴収する仕組みとした場合、当該受験者が置かれた経済的背景が受験できる回数に影響するおそれがある。また、試験場が近くに設置されている受験者にとっては複数回受験する際の移動・宿泊に係る物理的・経済的負担が小さい一方、離島・へき地在住の受験者をはじめ居住地の近くに試験場がない受験者にとってはそのような負担が大きいという、地域間格差が生じる可能性も想定される。一人の受験者による複数回受験について受験者や保護者を含む社会全体の理解を得るためには、このような格差の問題について検討する必要がある。
- また、複数回受験を認める場合、PBT でも起こり得ることだが、テストの点数を上げるための小手先のテクニック（test-wiseness）や何度も受験をして試験慣れをしたことにより高得点を取るケース、繰り返し受験する中でたまたま高得点を取るケースなどが生じることも想定される。
- さらに、複数回受験を認める場合には、各試験場が受け入れる必要のある「延べ受験者数」自体が大幅に増加することが想定されるため、それに対応できるよう試験日や座席数を確保する必要があることにも留意が必要である。

【コラム⑮】複数回受験を認める場合に生じる公平性の問題への対応

現行の共通テストの試験場は、原則として都道府県単位で大学が設定し、大学入試センターが志願者の分布や使用施設の収容数を考慮し指定している。現状では、約 700 の試験場のうち約 10 試験場を離島に設定しているが、それでもなお、試験場への移動に大きな負担を伴う受験者も少なくない。

共通テストで複数回受験を認める場合、試験場へのアクセスが容易な受験者と困難な受験者との間での公平性の問題が、これまで以上に拡大することが懸念される。特に、テストセンターを試験場として複数回受験を認める試験を実施する場合、その多くが都市部に立地していることに

留意する必要がある。

この問題を解消するためには、例えば、以下のような方策をとることが考えられる。なお、自宅・高校など受験者が自身で選択した場所で CBT 受験するという方策も考えられるが、本人確認・不正防止をどのように行うかが別途課題となる⁵⁸ため、ここでの検討からは除外する。

○試験場を離島・へき地を含む地方部にも設置する

離島・へき地を含む地方部に試験場を設置するためには、高等学校などを活用することが考えられる。そのためには、

- ・パソコンの調達（輸送による持ち込み又は現地のパソコンを活用）
- ・サーバ、ネットワークの整備及び環境確認
- ・その他の必要機器の整備

が必要となる。また、試験場に配置するスタッフを確保する必要があるが、PBT と比較すると、機器の設置や操作への習熟など、スタッフにはより多くのことが求められると考えられる。

○受験可能な回数に上限を設ける

複数回受験を認める場合でも、一人の受験者が受験可能な回数に上限を設けることで、地域間格差の拡大を抑えられる。なお、無制限に受験を認めると、経済的格差の問題や、試験場の座席数の不足の問題を引き起こす可能性があるが、受験可能な回数に上限を設けることでこれらの問題にも対処できると考えられる。

一方、受験可能な回数に制限を設けることは、目標の得点に届くまで何度も受験したい、模試代わりに受験したいなど多くの回数を受けることを希望する受験者から受験する機会を奪うことにもなりかねないことから、慎重な判断が必要である。

(3) 受験者の能力の経時的な変化の把握

- センター試験や現行の共通テストでは、異なる年度の試験結果を比較することは不可能だった。しかし、IRT に基づいて試験を実施することで、中長期的に試験結果を比較できるようになる。仮に、同一受験者が複数年にわたって受験した試験の結果から、当該受験者の能力推移を経年で比較することができるようになれば、例えば、ある受験者の大学入学前と大学入学後の能力を比較できるようになるなど、試験の結果をこれまで以上に様々な形で利用できるようになる可能性がある。また、異なる年度の受験者集団の能力推移の経年比較も可能になることから、さらに、高等学校・大学の教育改善や教育政策の検証・改善につなげることも考えられる。

⁵⁸ 自宅・高等学校など受験者が自身で選択した場所で CBT 受験する場合の本人確認・不正防止対策については、【コラム⑤】で詳述。

【コラム⑯】PISA 調査における得点の考え方～経年比較の観点～

OECD（経済協力開発機構）は、義務教育修了段階の15歳の生徒を対象に、読解力、数学的リテラシー、科学的リテラシーの三分野における学習到達度を国際的に調査する「生徒の学習到達度調査（Programme for International Student Assessment）」（以下「PISA 調査」という。）を実施している。PISA 調査はIRTに基づいて実施されている。2000年調査から2012年調査まではPBTで実施されていたが、2015年調査からCBTに移行し、直近の2018年調査においては、三分野のうち読解力についてアダプティブ方式⁵⁹で実施された。

PISAの主な調査目的は、その国の教育制度の長所や短所を明らかにし、政策立案に資する基礎的データを提供することにある。このため、PISA調査の結果は、生徒一人一人についてではなく、参加国・地域の生徒全体の平均得点により示される。IRTに基づいていることから、この調査結果は参加国・地域間で比較できる。【表コラム⑯－1】は、2018年調査における37のOECD加盟国における三分野の平均得点を、平均得点の高い順に上から並べたものである。

【表コラム⑯－1】PISA2018の結果（OECD加盟国（37か国）における比較）

OECD加盟国(37か国)における比較

は日本の平均得点と統計的な有意差がない国

	読解力	平均 得点	数学的リテラシー	平均 得点	科学的リテラシー	平均 得点
1	エストニア	523	日本	527	エストニア	530
2	カナダ	520	韓国	526	日本	529
3	フィンランド	520	エストニア	523	フィンランド	522
4	アイルランド	518	オランダ	519	韓国	519
5	韓国	514	ポーランド	516	カナダ	518
6	ポーランド	512	スイス	515	ポーランド	511
7	スウェーデン	506	カナダ	512	ニュージーランド	508
8	ニュージーランド	506	デンマーク	509	スロベニア	507
9	アメリカ	505	スロベニア	509	イギリス	505
10	イギリス	504	ベルギー	508	オランダ	503
11	日本	504	フィンランド	507	ドイツ	503
12	オーストラリア	503	スウェーデン	502	オーストラリア	503
13	デンマーク	501	イギリス	502	アメリカ	502
14	ノルウェー	499	ノルウェー	501	スウェーデン	499
15	ドイツ	498	ドイツ	500	ベルギー	499
16	スロベニア	495	アイルランド	500	チェコ	497
17	ベルギー	493	チェコ	499	アイルランド	496
18	フランス	493	オーストリア	499	スイス	495
19	ポルトガル	492	ラトビア	496	フランス	493
20	チェコ	490	フランス	495	デンマーク	493
	OECD平均	487	OECD平均	489	OECD平均	489
	信頼区間※(日本): 499-509		信頼区間(日本): 522-532		信頼区間(日本): 524-534	

※信頼区間は調査対象者となる生徒全員（母集団）の平均値が存在すると考えられる得点の幅を表す。

PISA調査は標本調査であるため、一定の幅をもって平均値を考える必要がある。

また、PISA 調査は調査年ごとに対象集団が違うが、継続して出題している問題（アンカー問題（anchor test））の正誤等の解答情報を活用することで、各国・地域の平均得点を経年比較できるようになっている⁶⁰。こうしたアンカー問題の正誤等の解答情報を手掛かりとして、異なる調

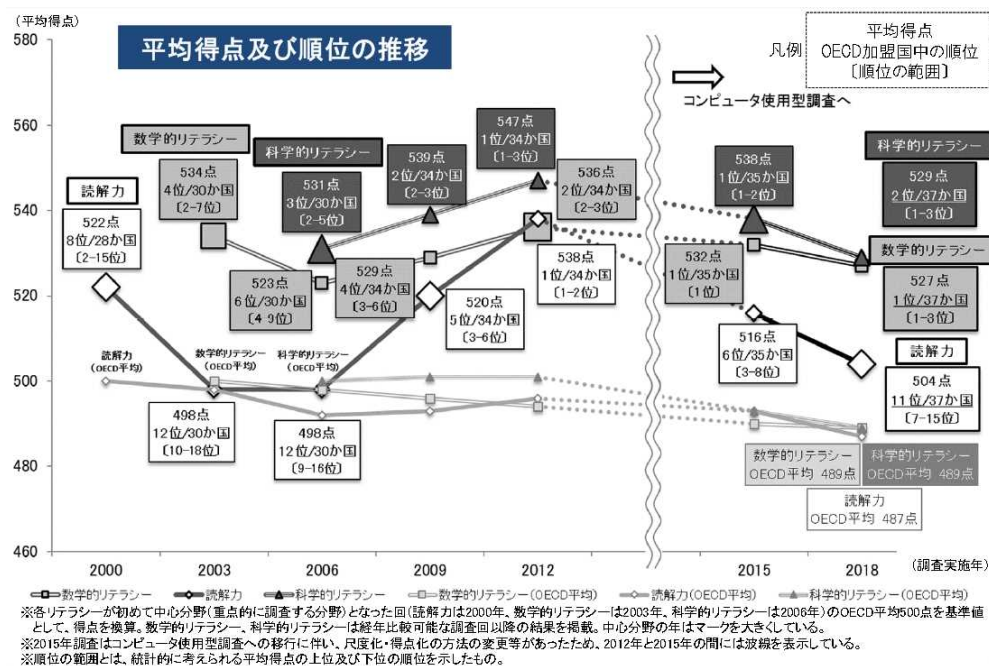
⁵⁹ アダプティブ方式については1. (1)で詳述。

⁶⁰ PISA 調査で出題される問題は、大きくアンカー問題と新規問題とに分けられる。例えば、2018年調査の読解力では、全245小問のうち72小問がアンカー問題である。72小問のうち、2000年調査から出題されているアンカー問題が28小問、2009年調査から出題されているアンカー問題が44小問である。

査年における比較をより正確なものにしている。

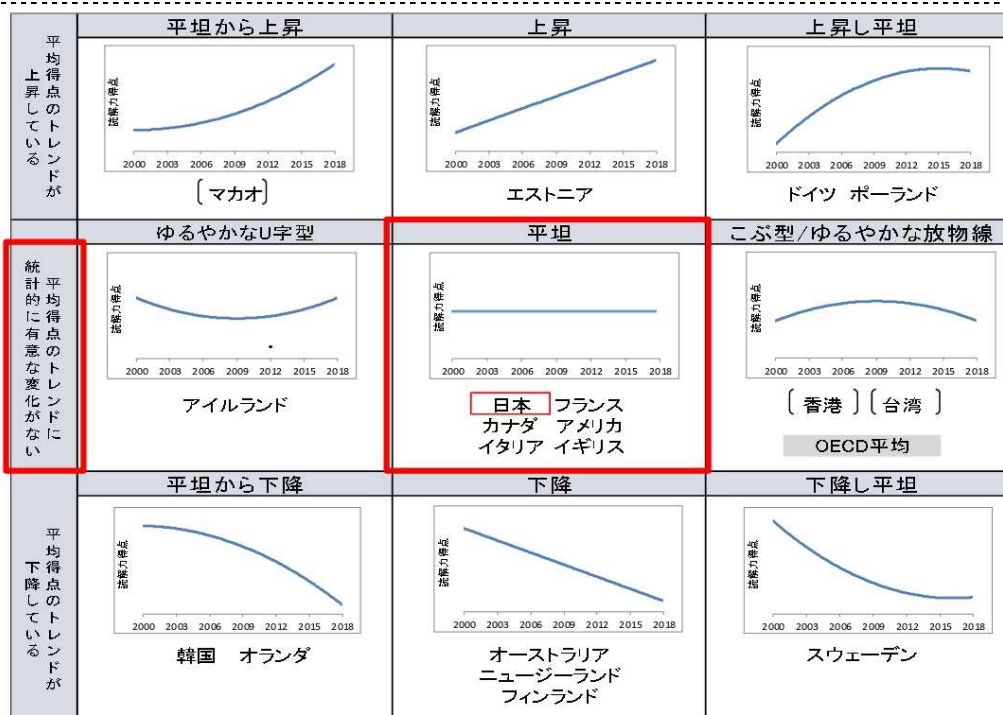
平均得点の表示も、各分野で得点の基準年を設定し⁶¹、その後の調査の平均得点を基準年のものに合わせることで、経年での比較が可能になっている。例えば、読解力は、最初に中心分野になった2000年調査時のOECD加盟国の平均得点を500点、標準偏差100点となるように測定単位となる得点の基準が定められた。2003年以降の調査においては、その基準と比較できるように統計処理し、得点化している。

実際に経年比較を行った資料を紹介する。2000年に始まったPISA調査は3年ごとに実施され、2018年調査で7回の調査データが蓄積されているが、【図コラム⑬-1】は2000年から2018年までの日本の平均得点と順位の推移を示したものである。また、OECDは、こうしたデータの蓄積により中長期的な変化の傾向も公表するようになってきている。【図コラム⑬-2】は、読解力について2000年から2018年の7時点の各国の平均得点に統計モデルをあてはめ、九つのパターンに分類したものである。



【図コラム⑬-1】PISA調査 日本の平均得点及び順位の推移

⁶¹ 三分野とも、最初に中心分野として調査が実施された年が基準年として設定されている（読解力：2000年、数学的リテラシー：2003年、科学的リテラシー：2006年）。



※〔 〕内は非OECD加盟国・地域

【図コラム⑬-2】PISA 調査 各国・地域の平均得点の長期トレンド（読解力）

ただ、PISA 調査の得点の経年変化を見る際の技術的な課題もある。例えば、PISA 調査は IRT に基づいて実施されているが、CBT が導入された 2015 年より前に開発された問題は 1 パラメタ・ロジスティックモデル、2015 年以降に新規開発された問題は 2 パラメタ・ロジスティックモデルに基づいて設計されている。このように問題によって基づく IRT のモデルが異なることにより、統計処理を施したとしても、平均得点を経年比較する際に精度が低下する可能性がある指摘されている。PISA 調査の平均得点の経年比較に当たっては、このような調査の設計や実施方法の変化にも留意する必要がある。

<出典>

文部科学省・国立教育政策研究所「OECD 生徒の学習到達度調査 2018 年調査（PISA2018）のポイント」（令和元年 12 月 3 日）

（執筆：大塚尚子（国立教育政策研究所国際研究・協力部総括研究官））

まとめ

1. 共通テストへの CBT や IRT の導入について

- 本報告書においては、大規模でハイステークスな使われ方をしている試験である共通テストに CBT や IRT を導入する場合、どのような利点、優れた点（メリット）や課題があるかについて、論点ごとに整理した。
- 大学入学者選抜、とりわけ共通テストでは、単なる学力試験・調査等をはるかに超える実施水準が求められる。現に、これまでの共通1次やセンター試験は、ミスやトラブルが非常に少ない安定的な実施がなされてきており、この枠組みを引き継ぐ共通テストにおいても、このような安定的実施が我が国の「試験文化」として期待されている。
- PBT で実施してきたセンター試験及び現行の共通テストの課題や CBT で実施した場合のメリットは大きいですが、現行の共通テストを CBT で行うこと、更には IRT に基づいて行うためには、第3章及び第4章で列挙した数々の課題を高いレベルで克服する必要がある。
- 具体的には、
 - ・全国的に均質で質の高い受験環境（パソコンやネットワーク等）の確保
 - ・トラブル等が生じた場合の対応体制の構築
 - ・新しい試験の在り方に対する受験者や保護者を含む社会全体の理解などについて、細やかな検討が必要と言える。

2. 今後求められる取組について

- これまで見てきたように、大規模でハイステークスな使われ方をしている共通テストを直ちに CBT 化することには、多くの課題があると言わざるを得ない。一方、小規模な試験、あるいは受験者や関係者に重大な結果をもたらすような使われ方をしないローステークスな試験としての活用であれば、CBT や IRT との親和性は高いと考えられる。
- また、本報告書は、大学入試センターの立場から、現行の共通テストへの CBT の導入に対象を絞ってまとめているが、本来的には単に PBT での実施を CBT での実施に切り替えることのみには焦点を当てるのは適切ではない。共通テストへの CBT の導入を検討する際には、今後次代を担う若い世代はもちろん、大学等での学び直しを希望する社会人を含む大学入学希望者に対して、どのような選抜を実現すべきなのか、その実現に CBT はどのように寄与するのか、という大局的な視点で、共通テストの性格や位置付けも含めて見直すことが必要と考える。

- これまで検討してきた CBT を導入していく上での課題, 長所や短所などを把握した上で, CBT の導入自体を目的化することなく, CBT を導入することの本来の意義を十分に引き出しながら, 受験者や保護者を含む社会全体が納得できる形を模索して, 国内外の最新の動向も踏まえつつ, 引き続き調査研究に取り組んでいくことが重要である。

※ 本報告書は令和 3 年 3 月時点の検討状況をまとめたものである。今後, 技術の進展や調査研究の進捗等に伴い, 内容に変更が生じる可能性がある。

国内外におけるCBTの実施例

参考資料 1

(1) 国内の資格検定試験等

試験名称	概要						
	試験場	年間受験者数	受験料	実施回数	試験時間	解答方法	結果発表
医療系大学間 共用試験	医学生・歯学生の能力と適正について全国的に一定水準を確保するための全国共通の標準試験。うち、臨床実習前の知識の総合的理解を問う試験がCBTで実施される。 【実施主体：(公社)医療系大学間共用試験実施評価機構】						
	47都道府県 /50以上(所属 大学で受験)	1 万人以上	25,000円	随時	360分	多肢選択	約10日後
英検S-CBT (注1)	英語のスピーキング、リスニング、リーディング、ライティングの試験。 【実施主体：(公財)日本英語検定協会】						
	47都道府県/ 約100	約400万人 (注2)	4,500 -12,600円	原則毎週土日	90-135分	音声吹込み・ 多肢選択・ 記述	約3-4週間後
SPI (注3)	企業が採用に利用する適性検査で、能力検査(言語・非言語)及び性格検査が出題される。 【実施主体：(株)リクルートマネジメントソリューションズ】						
	47都道府県/ 約50	200万人以上 (注4)	5,500 -6,500円 (企業負担)	随時	約70分	多肢選択	即時採点
ITパスポート 試験	情報処理技術者試験の一試験区分であり、「情報処理の促進に関する法律」に基づく国家試験。 【実施主体：(独)情報処理推進機構】						
	47都道府県/ 約120	10万人以上	5,700円	随時	120分	多肢選択	2-3時間後 (後日合格発表)
損害保険代理 店試験	損害保険を代理店として販売するための資格試験。 【実施主体：(一社)日本損害保険協会】						
	47都道府県 /約190	約100万人	1,900 -3,900円	随時	40-120分	多肢選択	3営業日後の 翌日

(注1) 2021年度4月以降に実施される「英検CBT」と「英検2020 1day S-CBT」を統合した方式について記載。

(注2) 実用英語技能検定、英検IBA、英検Jr. の2019年度志願者数の合算。

(注3) SPI3-U又はSPI3-UEをテストセンターで受験する方法について記載。

(注4) SPI3の全実施方法(テストセンター、インハウスCBT、WEBテスト、ペーパーテスト)の2019年度実績の合算。

(2) 国際的な資格検定試験等

試験名称	概要						
	試験場	年間受験者数	受験料	実施回数	試験時間	解答方法	結果発表
TOEFL iBT® テスト	主に大学・大学院レベルのアカデミックな場面で必要とされる、英語運用能力を測定する試験。 【実施主体:アメリカ合衆国, ETS(Educational Testing Service)】						
	[世界]150か 国以上 [日本]27都道 府県/約70	非公開	245USD (約26,000円) (注6)	月3-6回 年45回以上	約180分	音声吹込み・ 多肢選択・ 記述	約6日後
TOEIC®S&W	英語のスピーキング、ライティングの試験。 【実施主体:アメリカ合衆国, ETS(Educational Testing Service)】						
	12都道府県 /約40	3 万人以上	10,450円	年16回	80分	音声吹込み・ 記述	35日以内
ISAT	大学の医学部、歯学部、作業療法、スピーチ・言語セラピーコースに出願する留学生を対象とした、高等教育レベルでの学業成績の基礎となる知識と能力を評価する試験。批判的思考能力(Critical Reasoning)と数的能力(Quantitative Reasoning)が出題される。 【実施主体：オーストラリア, ACER(Australian Council for Educational Research)】						
	[世界]ー [日本]2都府 /2	ー	320USD (約34,000円) (注6)	年4回の試験期間 (1試験期間 = 1-2週間)	約180分	多肢選択	約2週間
GRE (注5)	大学院やビジネススクールに出願する受験生を対象とした試験で、分析論述(Analytical Writing)、言語能力(Verbal Reasoning)、数的能力(Quantitative Reasoning)が出題される。 【実施主体:アメリカ合衆国, ETS(Educational Testing Service)】						
	[世界]160か 国以上 [日本]2都県 /2	約65万人	205USD (約22,000円) (注6)	随時	約225分	多肢選択 記述	約10-15日後

(注5) GRE General Testについて記載。

(注6) 1USD=106円として計算。

(試験実施主体等の公表情報(令和3年2月19日確認)をもとに大学入試センター作成)

共通テストへの CBT 導入に係る検討を求める政府決定・報告等

教育再生実行会議「高等学校教育と大学教育との接続・大学入学者選抜の在り方について」（第四次提言）（平成 25 年 10 月 31 日）（抄）

3. 大学入学者選抜を、能力・意欲・適性を多面的・総合的に評価・判定するものに転換するとともに、高等学校教育と大学教育の連携を強力に進める。

（1）大学教育を受けるために必要な能力判定のための新たな試験（達成度テスト Z（発展レベル）（仮称））の導入

○ 達成度テスト（発展レベル）（仮称）は、その結果をレベルに応じて段階別に示すことや、各大学において多面的な入学者選抜を実施する際の基礎資格として利用することなど、知識偏重の 1 点刻みの選抜から脱却できるよう利用の仕方を工夫する。将来的には、試験問題データを集積し CBT 方式で実施することや、言語運用能力、数理論理力・分析力、問題解決能力等を測る問題の開発も検討する。

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改革について」（平成 26 年 12 月 22 日）（抄）

2. 新しい時代にふさわしい高大接続の実現に向けた改革の方向性

（1）各大学のアドミッション・ポリシーに基づく、大学入学希望者の多様性を踏まえた「公正」な選抜の観点に立った大学入学者選抜の確立

②入学希望者に求められる学力を評価する新テストの導入

「大学入学希望者学力評価テスト（仮称）」の在り方

◆ CBT 方式での実施を前提に、出題・解答方式の開発や、実施回数の検討等を行う。

高大接続システム改革会議「最終報告」（平成 28 年 3 月 31 日）（抄）

（3）「大学入学希望者学力評価テスト（仮称）」の導入

ウ 具体的な仕組み

⑥ CBT の導入

○ 思考力・判断力・表現力を構成する諸能力をテストによって評価するには、CBT の導入が有効であると考えられる。例えば、複雑な文章の構成力を問う問題や統計的方法を用いて複雑な現象を表現する問題の導入、多様な表現形態による様々な資料や動画等を活用した出題内容の拡大、テキスト入力等を利用した記述式問題の導入、音声入力を利用したスピーキングの評価、答えが一つに定まらない判断を相当回数伴う問題の導入、同一テスト時間内において問題の正答率に応じてそれ以降の問題の難易度を変えたりすることのできる適応型テストへの拡張、その他多くの展開を想定できる。

- 他方、C B Tについては、実施のための環境整備に時間を要すること、入学者選抜に係る大規模なテストにおける実施事例がないことなどに鑑み、導入には十分な準備が必要である。
- このため、平成36年度から始まると想定される次期学習指導要領の下でのテストからC B Tを実施することとし、現行学習指導要領の下での平成32～35年度間については、C B Tの試行に取り組む。試行においては、C B Tの導入に向けた実証的な検討の成果や、「高等学校基礎学力テスト（仮称）」の検討状況や実績などを踏まえつつ、端末の整備、システムの安定性・セキュリティの確保、機器導入・維持管理のコスト、その他の本格的実施に当たって前提となる課題について、専門家や民間事業者等の意見も聴きつつ十分な検討を行う。

「大学入学共通テスト実施方針」（平成30年7月13日文科科学省）（抄）

11. その他

- ※ C B Tの導入については、引き続きセンターにおいて、導入に向けた調査・検証を行う。平成29年度については、問題素案の集積方法の検討及び集積等を行う。この成果も踏まえ、平成36年度以降の複数回実施の実現可能性を検討する。

「未来投資戦略2018—「Society 5.0」「データ駆動型社会」への変革—」（平成30年6月15日閣議決定）（抄）

第2 具体的施策 II. 経済構造革新への基盤づくり [1]データ駆動型社会の共通インフラの整備

2. A I時代に対応した人材育成と最適活用 2-1. A I時代に求められる人材の育成・活用

(3) 新たに講ずべき具体的施策

i) 大学等におけるAI人材供給の拡大

- ・ 大学入学共通テストにおいて、平成36年度から必修科目「情報I」などの新学習指導要領に対応した出題科目とすることについて本年度中に検討を開始し、早期に方向性を示すとともに、コンピュータ上で実施する試験（CBT）などの試験の実施方法等について検討を進める。

「AI 戦略 2019」（令和元年 6 月 11 日統合イノベーション戦略推進会議決定）（抄）

II 未来への基盤作り 教育改革と研究開発体制の再構築

II－1 教育改革 （1）リテラシー教育

【大学入試・就職】

- ・ 大学入学共通テスト「情報 I」を 2024 年度より出題することについて CBT 活用を含めた検討

「成長戦略フォローアップ」（令和 2 年 7 月 17 日閣議決定）（抄）

1. 新しい働き方の定着 （2）新たに講ずべき具体的施策

xii) 大学等における Society5.0 時代に向けた人材育成

- ・ Society5.0 時代に必要な思考力・判断力・表現力などの学力を評価する大学入学共通テストを着実に実施していく。また、当該テストにおいて「情報 I」を 2024 年度から出題することについて CBT 活用を含めた検討を行う。

インドの国家試験の場合（一例）



試験場	テストセンター ・試験場数：約270か所(うち国外4か所) ・立地：インド国内の全エリアの受験者が12時間以内の移動でアクセスできるよう配置
実施方式	LAN方式
受験者数	約86万人
受験料	約1,200～13,000円 ※出願者のカテゴリー、会場によって異なる
試験日時	2月上旬の4日間、8つのスロットで実施(各日午前・午後の計2回)
試験問題	・全部で24科目 ・各科目とも「一般的な能力」と「専門分野」の2分野から構成(全65問) ・出題形式：テキスト、画像(グラフ等)、数式 ・解答方法：多肢選択又は数値入力

試験実施主体



試験時はLANを使用



ハードウェア（パソコンやネットワーク等）の整備

- ・テストセンターは「直営」と「非直営」の2種類。直営会場では建物に備わっているLAN環境を利用。非直営会場では、一時的に借用した会場にLAN環境を構築し、試験後に解体（元の状態に戻す）。
- ・CBT用途に開発されたOSが搭載されている、14インチ程度のディスプレイとウェブカメラが備えられたデスクトップ型のパソコンを使用。

トラブルへの対応

- ・以下のスタッフを配置。

会場管理者（1名/テストセンター）	会場管理責任者。試験会場レベルの全ての問題を管理。
ネットワーク管理者（2名/テストセンター）	会場のハードウェア（パソコンやネットワーク等）の正常動作について責任を持つ。
IT管理者（1名/250受験者端末）	

- ・トラブル時に備え、予備電源、セカンダリーサーバ、モバイルネットワーク、予備端末等を準備。
- ・予定どおり日時・場所での試験実施が困難となった場合は、予備日に実施。

※99.7%はトラブルなく実施されてきた。

身体検査のイメージ



不正防止策

- ・以下のスタッフを配置。

試験監督者（1人/20受験者）	受験者への案内・指示、試験中の監督等。
サポートスタッフ（1人/75受験者）	試験監督者らスタッフのサポート、受験者の案内・誘導。
セキュリティ管理者（1人/30受験者）	受験者入場時に顔写真・指紋画像の取得、再入場時の照合作業。
身体検査担当者（1人/80受験者、性別ごと）	身体検査を実施して携帯電話など不正な持込み等がないか確認。

- ・機器を使用した不正防止策も講じられている。

監視カメラによる監視イメージ



(例)試験中にウェブカメラで受験者の顔写真を撮影し、顔認証により出願時の顔写真と比較。
監視カメラによる監視。

障害等のある受験者への配慮

- ・視覚障害のある受験者の場合、追加費用でスクライブ（筆記者）サービスを提供。
- ・車椅子を使用する受験者は、車椅子で移動・利用可能なトイレを備えた1階又は移動可能な部屋の座席に割り当てられる。

(Y社からの報告に基づき大学入試センターが作成)

医療系大学間共用試験の例

※公益社団法人 医療系大学間共用試験実施評価機構 (<http://www.cato.umin.jp/cbt.html>) に掲載されている「2009 共用試験『医学系 CBT 公開問題』」より抜粋。

単一多肢選択形式（5つの選択肢から解答する。）

例題10 B-1-(4)-9) 神経管の分化と脳、脊髄、視覚器、平衡聴覚器と自律神経系の形成過程を概説できる。

胎齢 9 週の胎児横断面を図（カラー表示）に示す。
矢印で示す構造はどれか。

- A. 側脳室
- B. 第 3 脳室
- C. 中脳水道
- D. 第 4 脳室
- E. 中心管



例題27 C-7-(1)-1) 各消化器官の位置、形態と血管分布を図示できる。

腹部造影 CT（白黒表示）を示す。
胆嚢はどれか。

- A. a
- B. b
- C. c
- D. d
- E. e



多選択肢連問形式

(1つのテーマについて2問出題。6つ以上の選択肢から解答する。)

病態EMI(多選択肢連問形式) 例題63-1

連問 (1/2)

テーマ：心臓の圧容積関係

56歳の男性。健康診断で初めて高血圧を指摘されて来院した。脈拍 72 / 分、整。血圧 160 / 110 mmHg。胸部エックス線撮影で心陰影の拡大はない。

この患者があてはまるのは図（白黒表示）の①～⑨のどれか。

- A. ①
- B. ②
- C. ③
- D. ④
- E. ⑤
- F. ⑥
- G. ⑦
- H. ⑧
- I. ⑨

後負荷	高値	①	②	③
	正常	④	⑤	⑥
	低値	⑦	⑧	⑨
		低値	正常	高値
		前負荷		

病態EMI(多選択肢連問形式) 例題63-2

連問 (2/2)

テーマ：心臓の圧容積関係

56歳の女性。顔面のむくみと排尿回数の減少を主訴に来院した。体重は1か月前は52 kgであったが、現在は56 kgである。脈拍 78 / 分、整。血圧 140 / 80 mmHg。

この患者があてはまるのは図（白黒表示）の①～⑨のどれか。

- A. ①
- B. ②
- C. ③
- D. ④
- E. ⑤
- F. ⑥
- G. ⑦
- H. ⑧
- I. ⑨

後負荷	高値	①	②	③
	正常	④	⑤	⑥
	低値	⑦	⑧	⑨
		低値	正常	高値
		前負荷		

症候EMI(多選択肢連問形式) 例題65-1

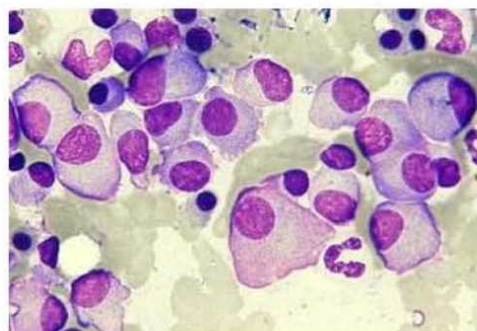
連問 (1/2)

テーマ：貧血

70 歳の男性。腰痛と全身倦怠感を主訴に来院した。家人に半年前から顔色が悪いと言われている。眼瞼結膜に貧血を認める。血液学所見：赤血球 350 万、Hb 10.7 g / dL、Ht 32 %、白血球 4,600 (桿状核好中球 10 %、分葉核好中球 49 %、好酸球 1 %、単球 2 %、リンパ球 38 %)、血小板 23 万。血液生化学所見：総蛋白 10.7 g / dL、アルブミン 4.1 g / dL。骨髓血塗抹 May - Giemsa 染色標本 (カラー表示) を示す。

考えられるのはどれか。

- A. 急性骨髄性白血病
- B. 急性リンパ性白血病
- C. 骨髄異形成症候群
- D. 再生不良性貧血
- E. 腎性貧血
- F. 多発性骨髄腫
- G. 鉄欠乏性貧血
- H. 慢性骨髄性白血病
- I. 慢性リンパ性白血病
- J. 溶血性貧血



症候EMI(多選択肢連問形式) 例題65-2

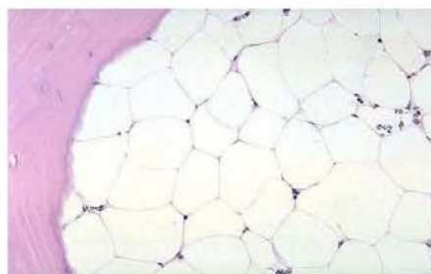
連問 (2/2)

テーマ：貧血

32 歳の女性。3 日前からの鼻出血を主訴に来院した。2 か月前から全身倦怠感を自覚している。皮膚は蒼白で紫斑と点状出血を認める。血液学所見：赤血球 215 万、Hb 7.0 g / dL、Ht 20 %、白血球 1,500 (桿状核好中球 1 %、分葉核好中球 15 %、好酸球 1 %、単球 2 %、リンパ球 81 %)、血小板 0.8 万。血液生化学所見：総蛋白 8.1 g / dL、アルブミン 4.1 g / dL。骨髓生検 hematoxylin - eosin 染色標本 (カラー表示) を示す。

考えられるのはどれか。

- A. 急性骨髄性白血病
- B. 急性リンパ性白血病
- C. 骨髄異形成症候群
- D. 再生不良性貧血
- E. 腎性貧血
- F. 多発性骨髄腫
- G. 鉄欠乏性貧血
- H. 慢性骨髄性白血病
- I. 慢性リンパ性白血病
- J. 溶血性貧血



出典情報 URL: <https://pathologyimages.com/cases/201/aml/aml-201-01.jpg>

順次解答 4 連問形式

(1つの症例について4問出題。次の問題に進むと前の問題に戻れない。)

順次解答4連問形式 例題70-1

連問 (1/4)

44歳の女性。6か月前から階段の昇降時や坂道を上るときに息切れを感じ、次第に強くなってきた。2か月前から動悸と疲労感が出現するようになり、顔色が青白いと家族にいわれて来院した。

医療面接として重要でない質問はどれか。

- A. 咳や痰が出やすいか。
- B. 生理の量が増えていないか。
- C. 下肢がむくむことはないか。
- D. 喉が渇きやすいことはないか。
- E. 歯肉から出血しやすいか。

順次解答4連問形式 例題70-2

連問 (2/4)

44歳の女性。6か月前から労作時の息切れ、動悸および疲労感が出現し、増悪したため来院した。咳や痰はないが、下肢がときどきむくむ。歯肉からの出血に気がついたことはない。偏食はなく、喫煙歴もない。6か月前から生理が不規則になり、月経2日目と3日目の月経血が多くなっている。

診られる可能性が高いのはどれか。

- A. 皮膚点状出血
- B. 頸静脈コマ音
- C. 頸部リンパ節腫脹
- D. 脾腫
- E. 深部感覚異常

順次解答4連問形式 例題70-3

連問 (3/4)

44 歳の女性。6 か月前から労作時の息切れ、動悸および疲労感が出現し、増悪したため来院した。半年前から月経の量が多くなっている。身長 155 cm、体重 49 kg。体温 36.2 °C。血圧 118 / 66 mmHg。顔面は蒼白で、眼瞼結膜に貧血を認める。頸静脈コマ音と収縮期心雑音を聴取する。腹部には異常所見を認めない。血液所見：赤血球 349 万、Hb 6.9 g / dL、Ht 22.4 %、MCV 64.1 fL、MCHC 30.8 %、白血球 4,000、血小板 26.9 万。血液生化学所見：総コレステロール 132 mg / dL、AST 8 IU / L、ALT 7 IU / L。末梢血塗末 Wright 染色標本（カラー表示）を示す。

診断のために必要な検査はどれか。



- A. セルロプラスミン
- B. ハプトグロビン
- C. ビタミン B₁₂
- D. フェリチン
- E. 補体価

順次解答4連問形式 例題70-4

連問 (4/4)

44 歳の女性。6 か月前から労作時の息切れ、動悸および疲労感が出現し、増悪したため来院した。半年前から月経の量が多くなっている。身長 155 cm、体重 49 kg。血圧 118 / 66 mmHg。顔面蒼白、眼瞼結膜貧血様。頸静脈コマ音と収縮期心雑音を聴取する。赤血球 349 万、Hb 6.9 g / dL、Ht 22.4 %、MCV 64.1 fL、MCHC 30.8 %、白血球 4,000、血小板 26.9 万。末梢血液塗末標本では赤血球大小不同と非薄赤血球が目立つ。血清フェリチン 0.5 ng / dL（基準 4.0 ~ 64.2）、総コレステロール 132 mg / dL、AST 8 IU / L、ALT 7 IU / L、Na 142 mEq / L、K 3.8 mEq / L、Fe 24 μg / dL。

この患者で血清鉄が低値となった機序はどれか。

- A. 鉄の吸収障害
- B. 鉄の利用能減少
- C. 鉄の還元能低下
- D. 鉄の体内貯蔵量減少
- E. 鉄の血清中輸送能低下

<付 録>

- 付録1 大規模入学者選抜における CBT 活用の可能性について（報告）（概要）・・・85
- 付録2 「問題バンク構築に係る調査研究について～CBT-IRT での共通テスト
「情報」の問題作成に係るフィージビリティの検証～（報告）」・・・・・・・・・・89

大学入学者選抜におけるCBT活用の可能性について（報告）

【概要】

大学入試センターでは、国内外の各種試験のCBT（Computer-based Testing）化の動向を踏まえ、平成23年から、研究開発部に所属する教育学やテスト理論等の教員によりCBTに関する研究を進めていたが、高大接続改革や情報教育の振興といった政府の施策の中でも、大学入学共通テストのCBT化の検討が求められた。

このため、研究開発部における調査研究に加え、外部委員による会議体として大学入学共通テスト企画委員会の下にCBT活用検討部会を設置するなど、テスト理論、情報・情報技術、初等中等教育、高等教育などの各分野の専門家の参画の下、共通テストにおけるCBT活用の可能性について検討を重ねてきた。

本報告書は、これまでの検討内容を取りまとめたものである。

【目次】

はじめに

第1章 PBTで実施する現行の共通テスト

第2章 共通テストにおいてCBTを活用する意義

第3章 パソコンやネットワーク等を活用した共通テストの実施

第4章 IRTに基づく共通テストの実施

まとめ

令和3年3月

独立行政法人 大学入試センター

1

はじめに

CBT（Computer-based Testing）とは

パソコンを用いて行われる試験全般を指す。従来行われてきたPBT（Paper-based Testing, 紙と鉛筆による試験）に代わり、近年、導入が進んでいる。

共通テストにおけるCBTの活用に関する提言

平成25年の教育再生実行会議第四次提言以降、共通テストへのCBTの導入について各所から提言。ただし、各提言の意図するところは必ずしも同じではなく、大きく以下の二つに分類できる。

- ①特定の科目に限らず**共通テスト全体にCBTを導入**する。
- ②共通テストにおいて**CBTを活用して「情報Ⅰ」を出題**する。

①共通テスト全体へのCBTの導入を提言

- 平成25年10月31日教育再生実行会議第四次提言
「高等学校教育と大学教育との接続・大学入学者選抜の在り方について」
 - 平成26年12月22日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改革について」
 - 平成28年3月31日高大接続システム改革会議
「最終報告」
 - 平成30年7月13日
大学入学共通テスト実施方針
- 上記では、IRT*等に基づく**複数回実施**についても併せて提言。

②CBTを活用しての「情報Ⅰ」の出題を提言

- 平成30年6月15日閣議決定
「未来投資戦略2018—「Society 5.0」「データ駆動型社会」への変革—」
- 令和元年6月11日統合イノベーション戦略推進会議決定
「AI戦略2019」
- 令和2年7月17日閣議決定
「成長戦略フォローアップ」

*IRT（Item Response Theory, 項目反応理論）：各受験者の問題に対する正答・誤答をもとに、問題の特性と受験者の能力を分けて推定する統計理論の一つ。

大学入試センターにおける調査研究

平成23年以降、教育学やテスト理論等の専門家によりCBTに関する調査研究を行い、さらに、テスト理論、情報・情報技術、初等中等教育、高等教育等の各分野の専門家を交え、共通テストにおけるCBTの活用について具体的に検討。2

第1章 PBTで実施する現行の共通テスト

第2章 共通テストにおいてCBTを活用する意義

現行の共通テストはPBTであるがゆえ、以下のような課題がある

出題・解答形式に制約がある

- 紙で表現できる形式の出題・解答のみ
- 取得できる解答情報はマークシートへのマークのみ（解答に至るまでの途中過程等の情報は取得不可）

問題冊子・解答用紙の印刷、輸送・保管、配付・回収が必要

- 50万人分の問題冊子等の印刷
- 倉庫で大学ごとに仕分けられ、輸送
- 各試験室で試験監督者が問題冊子等を受験者一人一人に配付
- 答案の返送に時間を要する（※1）とともに、マークシートの読取りによる解答の電子化が必要（※2）

試験問題は1バージョンのみで同一時刻一斉実施が必須

- 病気等の事情によって受験できなくなる（あるいは、受験はできても、試験問題の内容や当日の体調等に影響される）リスク
- 新型コロナウイルス感染症などの流行性疾患の感染拡大や大規模な自然災害発生があった場合、試験の実施自体が不可能になるおそれ

共通テストをCBTで実施すると…

パソコン上で出題・解答することで、多様なニーズに対応できる

- マルチメディア（動画、音声等）の利用など、多様な方法での出題や解答が可能
- 解答に至るまでの過程等の情報（操作ログ）の取得・活用

試験問題・解答を電子データにより配信・回収できる

- 受験者数の増減への柔軟な対応
- 問題訂正等への迅速な対応
- より効率的な採点の実現

共通テストをCBT-IRTで実施すると…

試験問題を複数バージョン用意して複数回実施ができる

- 試験日時の複数設定が可能
- 一人の受験生による複数回受験
- 受験者の能力の経時的な変化の把握

（※1）大学入学共通テストの規模の場合、答案等受領作業には約3日間を要する（交通状況や気象などの条件が悪い場合は回収困難に陥る可能性もある。）。

（※2）大学入学共通テストの規模の場合、マークシート式の解答用紙（約350万枚）を約4日間かけて2回ずつ読み取ることになる。

3

第3章 パソコン、ネットワーク等を活用した共通テストの実施

実施イメージ

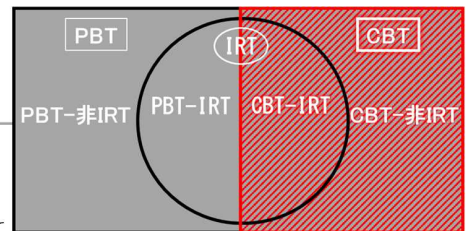
	既存のテストセンターで実施する場合	従来どおり大学等で実施する場合
	 テストセンター：CBTを受験するための試験会場	 ※現行の共通テストの会場を使用。
試験場	・立地：都市部中心 ・試験場の数：約180*	・立地：現行の共通テストの試験場と同様 ・試験場の数：全国約700
ハードウェア (パソコン、ネットワーク等)	・テストセンターのものを利用	・大学等で用意
試験実施に関わる業務 (試験場の準備、試験問題の保管・管理、試験監督等)	・主としてテストセンターが担当 ※1試験場当たりの座席数：10～200程度*	・主として大学等が担当 ※1試験場当たりの志願者数：最大4,000程度
トラブルへの対応	・予想外のトラブルに対応するため、予備のパソコンを用意、受験者の解答データを随時サーバにバックアップ ・テクニカル・スタッフを配置	
障害等のある受験者への配慮	・障害等がある受験者のアクセシビリティを考慮したCBT環境を整備 ・CBTでの受験が困難な受験生については、PBTでの実施も検討	

*国内大手テストセンター運営事業者X社の場合

4

パソコンやネットワーク等を活用した共通テストの 主な課題と必要な対応

※細字：課題，太字：必要な対応



ハードウェア (パソコン、ネット ワーク等)の整備

大規模なハードウェアの整備，及びその後の保管や保守が求められる。テストセンターや大学等に既に整備されているハードウェアを活用するのが合理的だが，その場合，パソコンの仕様やサイズ，ネットワークの回線速度などの条件が試験場によって異なることになる。

⇒**受験者や保護者を含む社会全体の理解が得られる形でハードウェアを整備するとともに，その後の保管や保守を確実にを行うことが必要。**

ソフトウェア (アプリケーション， ブラウザ等)の整備

現行のPBTで実施されている共通テストの特徴を生かしたままCBTを導入する場合，独自性の高い機能を備えたソフトウェアが求められる。

⇒**求められる機能が実装されたソフトウェアを，場合によっては独自開発により用意することが必要（独自開発する場合，開発後のクラウド環境の維持，改修，セキュリティ対応等も必要。）。**

試験実施時の トラブルへの対応

ハードウェアの保守を徹底するとともに，トラブル発生に備えて，予備のパソコンの用意や解答データのバックアップ，各試験場へのテクニカル・スタッフの配置等が求められる。それでもなおトラブルを皆無にすることが困難であることについて，受験者や保護者を含む社会全体の理解を得ることも重要。

⇒**トラブルの影響を除去できない場合，振替試験日の確保，代替の試験問題の準備等が必要。**

本人確認・ 不正防止策

PBTでの実施時には想定されなかった新しい形の不正が行われる可能性があり，生体認証（顔認証や指紋認証等）を活用した本人確認や，監視カメラ等の使用など，従来とは異なる新たな不正防止策が求められる。

⇒**個人情報やプライバシー等の保護の関係から，受験者や保護者を含む社会全体の理解を得ることが必要。**

経費

上述のような整備，対応等が求められるため，試験実施に必要な経費がPBTでの実施に比べて高額。

⇒**実施経費の増加に伴う財政負担について検討することが必要。**

※以下のような点にも留意が必要。

- ・一部の科目のみをCBTで実施するのはコストパフォーマンスが悪い。
- ・受験環境の整備やトラブルへの対応を考慮すると，同一時刻一斉実施ではなく分割実施（試験日時を複数設定）の方が実施しやすい。

第4章 IRTに基づく共通テストの実施

IRT（Item Response Theory，項目反応理論）とは

各受験者の問題に対する正答・誤答を基に，試験問題の特性と受験者の能力を分けて推定する統計理論の一つ
特徴①：異なる試験問題に解答した受験者同士の能力が比較できる（そのため試験の複数回実施が可能）
特徴②：統計的品質が管理された試験問題を出題できる

IRTに基づく試験のイメージ

問題作成

統計的品質管理を行った試験問題を大量（数千から数万問程度）に蓄積したデータベースである「問題バンク」を構築し，試験問題の非公開・再利用を前提に試験を実施することが多い。試験問題の曝露（試験問題が受験者の目に触れること）や漏洩への対応で，頻繁な問題入替え・追加が必要。

実施方式

（代表的な実施方式の例）

◆リニア方式

難易度を考慮しつつ，あらかじめセットされた試験問題で実施する方式。試験日や受験者によって解答する問題が異なる。



◆アダプティブ方式

1問ごとの正誤に応じて学力を測定し，正答すれば次に難しい問題が，誤答すれば次に易い問題が出題される方式。



成績の表示方法

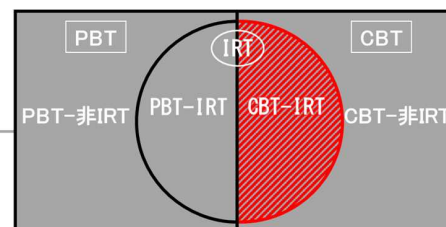
各設問に配点は設定せず，各設問の項目パラメータを用いたIRTの数式により受験者の能力値を推定し，それを基に得点を算出するという方法を採用することが多い。正答した問題の数と同じでも，難易度の高い問題に正答していると得点は高めに，難易度の低い問題に誤答していると得点は低めに算出される。

※難易度等のそろった等質な試験問題セットを使用する場合は，正答した設問の得点を足し上げた点数（素点）により示すことも可能。⁶

IRTに基づく共通テストの主な課題と必要な対応

パソコンやネットワーク等を活用した場合の主な課題に加えて、以下の課題等が生じる。

※細字：課題，太字：必要な対応



問題作成

作成すべき問題数が大幅に増える。試験の実施方法や受験者数によっては1科目数千～数万問の問題を作成することが求められる。同じ試験問題を本番の試験で何度も利用するという設計の場合、試験問題の曝露や漏洩への対応で頻繁な問題入替え・追加が必要。

⇒十分な数の問題作成者の確保，又は大学等の教員以外の人材を参画させる新たな問題作成の体制の構築が必要。

⇒実施経費の増加に伴う財政負担について検討することが必要。

試験問題の非公開

同じ試験問題を本番の試験で何度も利用するという設計とする場合，試験問題が原則非公開になる。

⇒得点の信頼性を維持するためには，試験問題の漏洩を防ぐことが必要（ただし，漏洩の影響をゼロにすることは不可能であり，そのことに対する理解も必要。）。

⇒（一部のサンプル問題を除いては）試験問題を教育現場で活用できなくなることへの理解が必要。

成績の表示方法

IRTの数式に基づいて算出した得点により成績を表示することが多いが，その場合，得点は現行の共通テストで用いられている「素点」とは異なる表示になる。

⇒「素点」とは異なる表示とすることとした場合，受験者自身が自己採点結果と得点の関係を理解するのが困難になるため，新しい成績表示方法について受験者や保護者を含む社会全体の理解を十分に得ることが必要。

上記の課題を克服できたとしても，試験の年複数回実施や一人の受験者による複数回受験の実現のためには以下の課題について検討が必要。

試験の実施時期

試験日を複数設定する場合，12月以前にも試験日を設定する必要が生じ得る。

⇒高等学校教育の実施に影響が出ないよう調整が必要。

⇒試験日が早いか遅いかによる不公平（感）を生じさせない工夫が必要。

複数回受験

受験生の経済的背景や居住地が複数回受験のしやすさに影響する。

⇒経済格差や地域間格差が生じないよう制度設計上の工夫が必要。

7

まとめ

➤ 大学入学者選抜，とりわけ大学入学共通テストでは，**単なる学力試験・調査等をはるかに超える実施水準が求められる。**

➤ 大学入学共通テストをCBTで行うメリットは大きい，そのためには先に列挙した数々の課題を高いレベルで克服する必要がある，

①全国的に均質で質の高い受験環境（パソコン，ネットワーク等）の確保

②トラブルが生じた場合の対応体制の構築

③新しい試験の在り方に対する受験者を含めた社会全体の理解

などについて，細やかな検討が必要である。

➤ CBTの導入自体を目的化することなく，CBTを導入することの本来の意義を十分に引き出しながら，受験者や保護者を含む社会全体が納得できる形を模索して，国内外の最新の動向も踏まえつつ，引き続き調査研究に取り組んでいく。その際，今後どのような大学入学者選抜を実現すべきなのか，その実現にCBTはどのように寄与するのか，という大局的な視点をもつことも重要。

問題バンク構築に係る調査研究について
～CBT-IRT での共通テスト「情報」の問題作成に係る
フィージビリティの検証～（報告）

独立行政法人大学入試センターにおいて、教科「情報」の CBT（Computer Based Testing）での実施に係る調査研究を行うため、平成 30 年度に「CBT 問題作成ワーキンググループ」（以下「問題作成 WG」という。）が設置された。問題作成 WG が設置された当初は、IRT（Item Response Theory, 項目反応理論）に基づく実施を前提とした問題作成は行っていなかったが¹、CBT は IRT との親和性が高く、共通テストを CBT で実施することを検討する際には、IRT に基づく実施についても視野に入れる必要があった。

IRT に基づいて試験を実施する場合には、あらかじめ識別力や難易度の値が付与されている大量の試験問題を蓄積した「問題バンク」を構築することも多い。そこで、問題作成 WG においては、令和元年度・2 年度に、問題バンク構築について、作問プロセス、作問体制、試験問題の曝露・漏洩を想定した類似試験問題の作成難度などに関して検討した。本報告書では、問題作成 WG が取り組んだ問題バンクに係る調査研究の内容や、そこから分かった課題等をまとめる。

¹ 平成 30 年度は、公募によって提供された問題素案を精選し、大学入学共通テスト（以下「共通テスト」という。）の試験問題作成の方向性に適合させながら知識・技能を基にした思考力・判断力等を測る CBT 形式の試験問題を作成した。作成した試験問題は IRT（Item Response Theory, 項目反応理論）のための「問題バンク」構築を前提とはせず、これまでの大学入試センター試験で問われているような大問形式の問題と、それぞれ独立した小問形式の問題を各領域で作成した。大問形式の問題は、全体として文脈がある中で各小問が連携する形で構成され、深い思考力・判断力等を評価することが可能となっている。小問形式の問題は IRT を念頭において作問し、単なる知識・技能を問う問題でなく、思考力・判断力等を評価できるか注意深く検討した。最終的に大問 5 問、小問 16 問を作成した。これらの問題の難易度や識別力などの項目パラメタを推定するために、予備調査として平成 31 年 2～3 月に高等学校（9 校 453 人）において実証実験を実施した。

【参考】CBT の特徴を生かした出題について

CBT を導入し、端末で試験を実施する場合、以下のような特徴をもつ問題を出題できるようになる。

- ・試験問題の多様性（例えば、動画を見たり、音声を聞いたりして解答する試験問題など、メディアを豊富に利用した試験問題）
- ・解答方法の多様性（画面に表示された図表の該当箇所を選択する、画面に表示された文章の該当箇所に下線を引く、文章を入力するなど、多様な形式の解答を要求する問題）
- ・受験者と試験問題の双方向性（受験者が解答した内容に対してフィードバックを行い、そのフィードバックを踏まえて受験者がさらに試行錯誤をして解答を導くというような問題）

特に、教科「情報」の場合、CBT で試験を実施することで、次のような出題を実現でき、これからの時代に必要な資質・能力を、より適切に測定できるようになる。

- ・プログラミングに関する試験問題において、受験者が作成したプログラムを実際に実行し、出力結果に応じてプログラムを修正して解答を確定させる（【参考資料 1】）。
- ・データの活用に関する試験問題において、実データをもとに項目やグラフの種類、集計値を選び、出力されたグラフやクロス集計の結果から問題の解答を導く。

1. 調査研究の目的

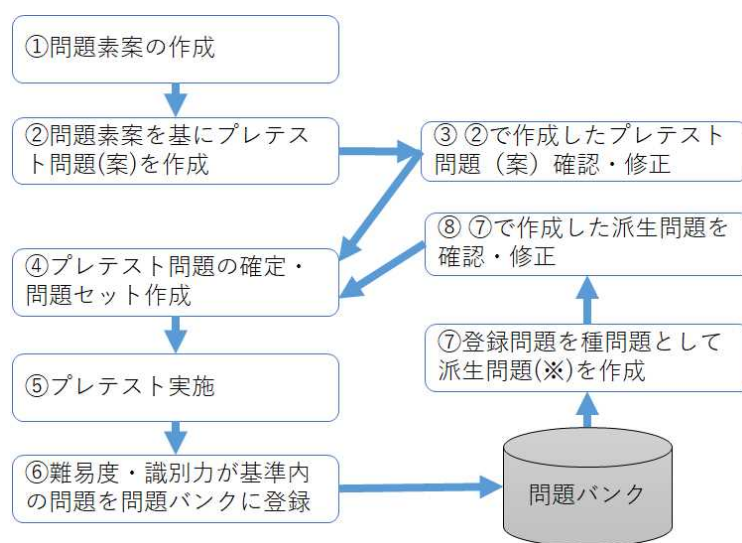
問題作成 WG では、共通テストにおいて教科「情報」の試験を CBT で実施する場合の問題作成について調査研究を行った。特に令和元年度からは、共通テストにおいて教科「情報」の試験を、問題バンクを構築した上で IRT に基づく CBT（以下「CBT-IRT」という。）で実施すると仮定した上で試行的に問題バンクを構築し、主に問題作成の生産性を中心にそのフィージビリティ（実現可能性）を検証する。

2. 問題作成 WG の取組内容

(1) 問題バンクを構築するプロセス・体制

問題バンクを構築する方式の CBT-IRT で試験を実施する場合、多数の試験問題、試験の特性や実施方法によっては数千～数万問の試験問題を用意する必要がある。試験問題を大量に用意するためには、従来とは異なる、新しい問題作成プロセスや問題作成体制を確立する必要がある。問題作成 WG では、共通テスト「情報」の問題バンクを構築する場合の問題作成プロセスや問題作成体制を具体的に想定し、そのプロセスや体制を小規模ながら実践し、試行的に問題バンクを構築することとした。

まず、問題作成プロセスについては、IRT に基づいて実施する他の試験²の問題バンク構築のプロセスも参考にし、【図1】のように進めることを想定した。

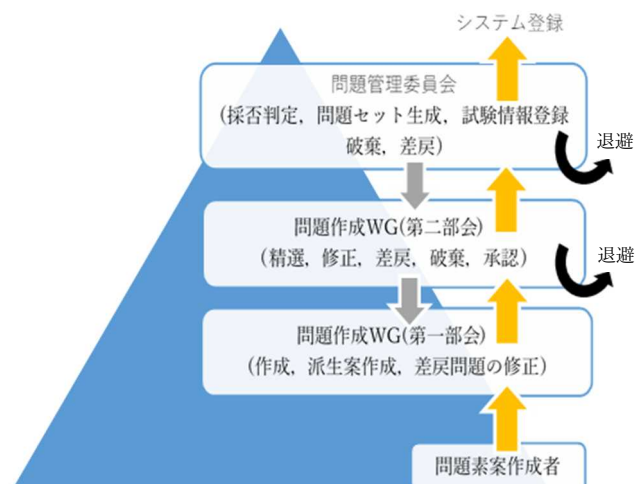


【図1】問題バンクを構築するプロセス

このプロセスに沿って問題作成を行うための問題作成体制については、継続的に大量の問題を作成し続ける必要があるため、【図2】のように複数の会議体を設置して役割を分担しながら行うことを想定した。

以上のような想定に基づき、問題作成 WG 委員約 20 人（高校・大学関係者等）を役割分担した上で、問題バンクに蓄積する試験問題の作成を行った。具体的には、問題作成 WG 委員を【表1】に示す5つの班に分け、1～4班が問題の作成を行う「第一部会」の役割を担い、5班が「問題管理委員会」の役割を担うこととした。「第一部会」では、1～4班の各班が科目「情報Ⅰ」の4つの領域ごとのうちそれぞれ1つを担当することとした。また、「第一部会」が作成した問題の確認・修正を行う「第二部会」については、1～4班の各班が、他班が「第一部会」で作成した問題の確認・修正を行う形でその役割を担うこととした。

² 公益社団法人医療系大学間共用試験実施評価機構が実施する、医療系大学間共用試験（臨床実習において患者の診療に参加する学生の知識、態度及び技能が標準的な水準に到達していることを評価する共通の評価試験）など。



【図 2】問題バンクを構築するための問題作成体制（イメージ）

【表 1】問題作成 WG 内の役割分担

	1 班	2 班	3 班	4 班	5 班
主担当領域	(1) 情報社会 の問題解決	(2) コミュニ ケーションと 情報デザイン	(3) コンピュ ータとプログ ラミング	(4) 情報通信ネ ットワークと データの活用	—
問題管理委員会	—	—	—	—	採否判定
問題作成 WG (第二部会)	他班が作成 した問題の 確認・修正	他班が作成 した問題の 確認・修正	他班が作成 した問題の 確認・修正	他班が作成 した問題の 確認・修正	—
問題作成 WG (第一部会)	主担当領域の問題作成				—

(2) 試行的な問題バンク構築

問題作成 WG は、令和元年度・2 年度において、(1)で説明したプロセス・体制により問題バンクの構築に試行的に取り組んだ。その取組内容を具体的に紹介する。

①問題素案を作成

大学入試センターにおいては、平成 30 年 7 月から 9 月にかけて、平成 30 年告示高等学校学習指導要領の教科「情報」の科目「情報 I」について、具体的な問題のアイデアを大学等の研究者や高等学校の教員等から広く募集した。この公募により、384 件の提供を受けた。提供された問題のアイデアから問題形式や題材についてヒントを得ながら、令和元年度に個々の問題作成 WG 委員が問題素案を作成した。作成された問題素案は 181 問であった。

②問題素案を基にプレテスト問題（案）を作成

「第一部会」の役割を担う 1～4 班が、①で各委員が作成した問題素案を基にしながら

ら、作題方針や測りたい資質・能力に合わせて、プレテスト問題（案）を作成した。

③②で作成したプレテスト問題（案）を確認・修正

②で作成されたプレテスト問題（案）を、「第二部会」の役割を担う 1～4 班が確認・修正した。必要に応じて「第一部会」において再度検討することとして差し戻すことも行った。「第二部会」が確認・修正・差し戻しをし、最終的に承認したものを「問題管理委員会」に送った。

④プレテスト問題の確定・問題セット作成

「問題管理委員会」の役割を担う 5 班が、「第二部会」が承認した後のプレテスト問題（案）の中からプレテストで出題する問題を選定した。令和元年度末までに、科目「情報Ⅰ」の 4 領域にわたって合計 98 問の問題が「問題管理委員会」によりプレテストでの出題に適したものとして承認された（【表 2】）。

ただし、令和元年度はプレテストの準備に要する時間の関係で、令和元年 11 月までに承認された 62 問のみ⑤のプレテストで使用した。大学入試センターにおいて、この 62 問について、限られた数の参加者からより多くのデータが得られるよう工夫し、プレテスト用の問題セット（8 種類）を作成した。

項目パラメタ（難易度や識別力等）をより正確に推定するためには、1 問あたり 200～500 人程度の参加者が必要とされるが、今回の実証実験では、1 問あたり約 160 人の参加者を割り当てるにとどまった。

⑤プレテストの実施

令和元年 11 月までに承認された 62 問の問題について、問題の項目パラメタを推定するために、令和 2 年 1～2 月に、④で作成した問題セット（8 種類）を用いてプレテストを想定した実証実験を実施し、参加者（高校生 307 人、大学生 340 人）に解答してもらい、その解答情報を取得した。

⑥問題の特性が基準内の問題を問題バンクに登録

⑤で出題した 62 問についての参加者の解答情報を基に、項目パラメタを推定した。具体的には、各問題の正答率や I-T 相関³、正答者平均解答時間を算出し、これら全てにおいて基準を満たした 32 問を問題バンクに登録した（【参考資料 2】）。

⑦登録問題を種問題として派生問題を作成

³ I-T(Item-Total)相関とは、各問の得点と合計得点との関係性を表す指標の一つ。

問題の識別力の指標として用いられる。合計得点の高い人がその問題に正解し、合計得点の低い人がその問題に誤答していた場合、I-T 相関は 1 に近いプラスの値をとる。

問題バンク構築には、より多くの問題を効率よく作成することが重要である。そこで、令和2年度は一つの問題（これを「種問題」と呼ぶ。）から、同様な形式の問題（これを「派生問題」と呼ぶ。）をどのような考え方でどの程度の問題数を作成することができるかを検証した。

まず、⑥で問題バンクに登録した様々な形式の問題32問から、それぞれどの程度派生問題を作成することができるかを実際に試した。ここでの派生問題は、種問題となる問題の形式は変えず、正答に至るプロセスも変わらないものとし、条件や状況設定を単純化したり複雑化したりすることによって、問題を効率良く作成できるものであり、種問題の形式や内容によって、派生問題を作りやすい問題とそうでない問題、派生問題の適正な作問数を検討してみた。

32問の種問題を基に派生問題の作成に取り組んだ結果、作成できる派生問題のタイプによって、32問の種問題を大きく次の三つに分類することができた⁴。なお、この分類は、あくまでも問題作成WGで作成した教科「情報」の問題に対して便宜上行ったものであり、妥当性や他教科への適応可能性などに関してさらなる検討が必要である。

【表2】種問題の分類とそこから作成可能な派生問題の特徴

種問題の分類	作成可能な派生問題の特徴	該当の種問題の数
種問題 (T ₁)	種問題と問題形式及び設定が同じであり、難易度もあまり変わらない (H ₁)	16 問
種問題 (T ₂)	種問題と問題形式及び設定が同じであるが、与える条件によって難易度を変えることができる (H ₂)	10 問
種問題 (T ₃)	種問題と問題形式は同じであるが、設定及び条件を変えることで難易度を変えることができる (H ₃)	6 問
合計		32 問

※各分類の具体的なイメージは【参考資料3】参照

⑧⑦で作成した派生問題を確認・修正

令和2年度は実際には⑧のプロセスは行っていないが、⑦で作成された派生問題は「第二部会」で確認され、「問題管理委員会」にてプレテストの問題として確定されることになる。ここでの作業は、問題としての内容や質の確認という側面よりも、種問題との関係性において適当な派生問題であるかを検討することになる。

①～⑦で小規模ながら試行的に実施した問題バンク構築プロセスにおいて作成された問題数の推移は【表3】のようになった。

・作成されたプレテスト問題（案）148題のうち、最終的に問題管理委員会でプレテス

⁴ 他の分類として統計データや社会情勢、技術進歩に影響される問題か否かなども考えられる。

トでの出題に適したものとして承認されたものは 98 題で約 66%であった。

- ・プレテストを想定した実証実験で出題された 62 題のうち、正答率や I-T 相関、正答者平均解答時間が基準範囲内のものは 32 題（約 52%）であった。

【表 3】 作成した問題数の推移

年度	令和元年度				令和 2 年度
ステータス	①	②③	④⑤	⑥	⑦⑧
	問題素案	プレテスト 問題（案）	プレテスト 問題 （※ 1）	問題バンク に登録 （種問題）	派生問題 （※ 2, 3）
(1) 情報社会の問題 解決	38	38	20 (8)	0	— (0,0,0)
(2) コミュニケーシ ョンと情報デザイン	35	31	21 (11)	7	— (4,2,1)
(3) コンピュータと プログラミング	58	29	26 (25)	16	— (6,8,2)
(4) 情報通信ネット ワークとデータの活 用	50	50	31 (18)	9	— (6,0,3)
合計	181	→ 148	→ 98 (62)	→ 32	— (16,10,6)

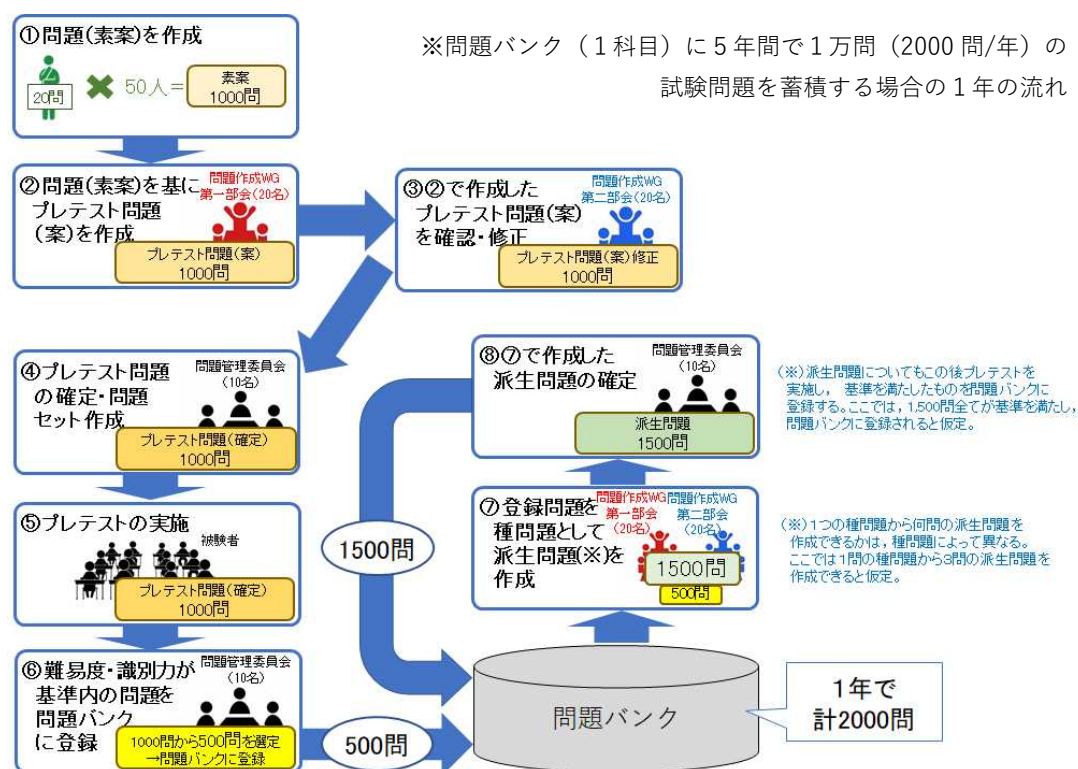
※ 1 （ ）内は令和元年度実証実験（プレテストを想定）の出題問題。

※ 2 （ ）内は種問題を T₁, T₂, T₃ に分類した際の内訳。

※ 3 ⑧（確認・修正）を行っていないため、派生問題数は「—」としている。

(3) 問題バンク構築に必要な人員・経費（試算）

(1)及び(2)で取り組んだ体制に準じて、問題バンクに蓄積する試験問題を 5 年間で 1 万問作成すると仮定した場合、(2)の最後で整理した作成できた問題数の実績を踏まえると、【図 3】のような流れにより 1 年間で 2,000 問の問題を作成することになる。



【図3】問題バンク構築の流れ（イメージ）

【図3】のように試験問題を作成する場合、1年あたり、次の人員・日数が必要となる⁵。

- ・素案作成者：50人（各作成者が20問作成，1000問の問題素案を作成。）
- ・第一部会：20人（会議回数は50回/年，問題（案）を作成。）
- ・第二部会：20人（会議回数は50回/年，第一部会が作成した問題（案）を確認・修正）
- ・問題管理委員会：10人（会議回数は25回/年，第二部会が確認・修正した問題（案）を承認）
- ・プレテスト参加者：20,000人⁶

また，上記の人員により問題作成を行うと仮定すると，1年あたり以下の経費を要することになる。

⁵ これ以外にも必要となる人員・日数は想定されるが，ここでは主だったものを取り上げることとする。

⁶ プレテスト参加者数は，

- ・1問あたり延べ200人の参加者が必要
 - ・参加者1人が25問を解答
- という仮定に基づいて算出した。

- ・素案作成：0.5 万円×20 問×50 人＝500 万円/年（素案 1 問につき 0.5 万円の謝金を支払うと仮定）
- ・第一部会：50 万円×50 回＝2,500 万円（会議費を 50 万円/回と仮定）
- ・第二部会：50 万円×50 回＝2,500 万円（会議費を 50 万円/回と仮定）
- ・問題管理委員会：25 万円×25 回＝625 万円（会議費を 25 万円/回と仮定）
- ・プレテスト：0.2 万円×20,000 人＝4,000 万円（参加者 1 人当たり 0.2 万円の謝金を支払うと仮定）

1 年当たり 1 億 125 万円

5 年間総額 5 億 625 万円

3. 共通テスト（情報）の問題バンク構築における課題

小規模ながら問題バンク構築の一連のプロセスを試行することによって、現行の共通テストの位置付けや問題作成方針を維持したまま、共通テスト（情報）を CBT-IRT で実施するために問題バンクを構築する場合、様々な課題があることが見えてきた。課題としては、主に次の 4 点があげられる。

(1) 共通テストとしての問題の質と量の両立について

・共通テストの問題作成方針に沿った試験問題の大量作成が必要

従来、大学入試センターでは、複数の小問を組み合わせた「大問形式」の出題を行っている。大問形式の場合、小問形式に比べて、試験問題の大量作成が難しい、試験問題の曝露・漏洩のリスクが高いなどの課題が懸念され、大問形式のまま IRT を導入する場合には、この課題について検討を要する。

一方、小問形式の場合、大問形式のように試験問題の展開の中で思考の過程を段階的に示すことが難しい。このため、共通テストの問題作成方針に沿った試験問題を大量に作成することは難しいのではないかと考える。

・1 問 1 問の問題作成に割ける時間が少ない中での試験問題の質の維持が必要

試験問題の大量作成が必要となるため、従来の作問プロセスより 1 問 1 問の問題作成に割くことができる時間が少なくなる。このため、短い検討時間においても質の高い試験問題を作成するための工夫が求められる⁷。

・派生問題の作成が容易ではない中での試験問題の大量作成が必要

令和 2 年度の問題作成 WG の派生問題作成の取組に限って言えば、共通テスト「情報」の試験問題を念頭に、種問題から派生させるという手法で作成するのは容易ではなかった。

⁷ なお、問題作成WGにおいて作成した試験問題の一部は、「平成 30 年告示高等学校学習指導要領に対応した大学入学共通テストの『情報』の試作問題（検討用イメージ）」において活用された（【参考資料 4】）。この試作問題は、情報処理学会、日本教育工学会、教育システム情報学会、日本産業技術教育学会、8 大学情報系研究科長会議などから高評価を得ている。

まず、2. (2)⑦で述べた種問題 $T_1 \cdot T_2$ 及び派生問題 $H_1 \cdot H_2$ は、種問題又は派生問題のいずれかが曝露・漏洩の影響を受けると、関係する試験問題全ての項目パラメタに大きな影響が出て、大幅な入替えが必要になってしまうことがある。種問題 $T_1 \cdot T_2$ から派生問題 $H_1 \cdot H_2$ を作成すること自体は可能だが、これらは種問題と問題形式や設定が同じであるため、同じ解法パターンで解答することになるので、いずれか一つが曝露・漏洩されて解法パターンが明らかになると、難易度が大幅に下がると予想される。

一方、種問題 T_3 及び派生問題 H_3 は、種問題又は派生問題の一部が曝露・漏洩されたとしても項目パラメタの変動は比較的小さいと考えられる問題である。このため、問題形式によってはある程度の数の派生問題の作成は可能だが、作成できる派生問題の数には限界がある。

このように、少なくとも共通テストの教科「情報」の問題バンクの構築及び維持するためには、特定の種問題から多数の試験問題を派生させて作成し、それを再利用していくという手法を取るのが難しいため、幅広い領域の中で異なる資質・能力を測ることができる大量の問題を作成していく必要がある。

(2) 問題作成体制について

・問題作成の人材も含めた新たな体制の構築が必要

これまでの大学入試センター試験については、50 万人以上が受験する大学入学者選抜として求められる正確性や質を確保するため、問題作成・点検に携わる現役大学教員を中心とする大学関係者の創造的なアイデアと入念な点検、絶え間ない修正のプロセスがあり、大学関係者の多大な時間と労力を割いて少数の試験問題を練り上げてきた⁸。しかし、問題バンクに蓄積する試験問題の作成全体を、現役大学教員の協力のみに頼り続けた場合、これまで以上の負担を強いることになる⁹。

問題バンクの構築を支えるためには新たな体制づくりが必要となるが、その方法としては、①アイテム・ライター（問題素案を作成し、それを試験問題として適切な形に整える者）を専属で大学入試センターに配置する、②問題作成を外注するなどが考えられる。いずれの場合でも、問題作成に関する細目(item specification)をあらかじめ示すなど、問題バンクに蓄積する試験問題に求められる条件を確実に共有する必要があると考える。

⁸ 情報漏洩防止の観点から、問題作成業務は大学入試センターの問題作成エリア内のみで行うため、問題作成者は年間 50 日程度大学入試センターに出張する（1 回の出張で 3～5 日間、終日センターで問題作成に従事）。また、問題の正確性や質を高めるため、問題完成までに複層的な体制での点検作業も実施している。

⁹ 令和元年度は 19 人の委員による会議を 9 日開催したが、問題バンクに登録できた問題は 32 問であった。数千～数万問の問題作成を実現するには、問題作成のための会議の委員数や会議の回数を大幅に増やす必要がある。

また、これまでの問題作成・点検の作業は、全国から大学入試センターに委員が集まり限られた日数の制約の中で作業が行われていたが、問題バンク構築及び維持管理には、年間を通して逐次問題作成及び点検の作業が必要になってくる。加えて、問題作成に係る多くの人材登用と機会の創出を考えた場合、従来からの方式を改め、オンラインによる問題作成システムでの問題作成・点検の可能性についても検討していく必要がある。この場合、良質な問題の創出に不可欠な協議が十分できるか、また、どこまでオンラインでの作業を認めるかも含め情報セキュリティ（特に機密情報の管理）の確保についても慎重に検討を重ねていく必要があるのは言うまでもない。

(3) 難易度・識別力の付与について

・事前調査（プレテスト等）に時間と経費を要することを踏まえた検討が必要

問題の難易度や識別力などの項目パラメタを推定するために、共通テストの受験者と同等な学力を持つと想定される参加者に、プレテスト等により事前に調査する必要がある。

項目パラメタをより正確に推定するためには、1問あたり200～500人程度の参加者が必要とされる。仮に1問あたり延べ200人の参加者が必要とすると、問題バンクに1万問を蓄積する場合、2. (3)の試算に基づくと、5年間総計で延べ10万人の参加者に対して事前調査を実施することが必要になるなど、試験問題の蓄積に時間と経費を要することを踏まえた検討が必要である。

なお、共通テストでは学習指導要領を踏まえて問題が作成されるが、難易度や識別力などの統計情報をより正確に把握するためには、事前調査で出題される試験問題に対応した学習指導要領の下で学んだ人を対象に事前調査を行うことが望ましい。そのような条件を付した場合、事前調査の参加者になりうるのは高校を卒業して間もない者及び現役高校生となるが、事前調査での参加者は試験問題の内容を知ることになるため、今後共通テストを受験する可能性のある現役高校生を参加者とすることは避けるのが適当で、高校卒業直後の者から多数の協力を得る必要があることにも留意が必要である。

(4) 問題バンクのメンテナンスについて

・特に共通テストの場合、頻繁な問題の入替が必要

試験を何回か実施する中で問題が曝露されたり漏洩したりすることにより難易度や識別力が変動するが、試験の測定精度を維持するためには、曝露・漏洩により項目パラメタが変動した問題を引退させ、新しい問題を問題バンクに入れることが必要になるのである。特に共通テストは、ハイステークスな使われ方をする試験であり、試験の精度を保つ必要性が高い反面、漏洩のリスクも高いため、新規問題の作成・問題バンクへの追加の必要性が他の試験以上に高いと思われる。

・統計データや社会情勢等の変動に伴う試験問題の修正が必要

試験問題に使用している統計データや社会情勢等に変動などがあつた際に、関係する

試験問題のメンテナンスが必要となる。そのために、試験問題ごとに適切にメタ情報を付ける必要があるが、大量の試験問題の中からメンテナンスが必要な問題を確実に確認するメタ情報付けは容易ではない。このため、例えば、問題文中のキーワードと思われるものを自動検出してメタ情報候補として提案するシステムを開発するなど、コンピュータの力も借りながら、統計データ等の変動に伴う修正が可能な体制を整えることが求められる。

・学習指導要領改訂への対応が必要

おおむね 10 年に一度学習指導要領が改訂される度に、問題バンク中の全ての試験問題の内容が新しい学習指導要領に対応しているか否かを確認する必要がある。対応していない試験問題が多い場合、大幅な入替作業が必要となる。

・テスト理論の専門家の配置が必要

曝露・漏洩により難易度や識別力といった項目パラメタが変動した試験問題を適正に問題バンクから除外するためには、項目パラメタの推移を的確に見極め、長期的にモニタリングができることができるテスト理論の専門家を大学入試センターに配置することが望ましい。

4. まとめ

このように、現行の共通テストの枠組みにおいて、教科「情報」の試験を、問題バンクを構築した上で、CBT-IRT で実施することを目指す場合、問題バンクの構築自体にも数々の課題があることが、問題作成 WG の試行的な取組から明らかになったと言える。さらに、問題作成 WG の 3 年間の活動ではできなかったが、CBT-IRT を前提に、小問を中心に出题することとした場合、共通テストで求められている思考力・判断力等を十分に評価できるのかを検証することが必要である。一方、教科「情報」を含め共通テストを CBT で実施するメリットがあることも事実である。共通テストへの CBT の活用については、問題バンクを用いない方法も含めて、引き続き検討を行っていく必要がある。

CBT の特徴を生かした出題の例（プログラミング問題）

CBT で試験を実施する場合、「情報」のプログラミングの問題については、プログラミングの実行環境が与えられることで、受験者が作成したプログラムを実際に動かし、出力結果に応じてプログラムを修正して解答を確定させるといったことも可能になる。

例えば、次ページは、配列の要素を昇順に並べ替える基本的なプログラミング（バブルソート）について PBT（マーク式）で出題する場合の問題例であるが、問題作成者がつくったプログラムを読み解くことが中心の出題となっている（ただし、この問題では、バブルソートの二通りの考え方いずれによって解答しても正答となるよう工夫している。）。

一方、同じ内容を CBT で出題する場合、受験者が自分でプログラムを組み立て実行することによって、トライ＆エラーの過程を経ながら、目標に近づけていくといった解き方をさせることが可能になる。また、複数の考え方を容認する実際のプログラミングに近い形での出題が可能になる。

PBT 問題 （マーク式）	・問題作成者がつくったプログラムを読み解くことが中心の出題となる。 ※ただし、次ページの問題例では、バブルソートの二通りの考え方いずれによって解答しても正答となるよう工夫している。
CBT 問題	・受験者が自分でプログラムを組み立て実行することによって、トライ＆エラーの過程を経ながら、目標に近づけていくといった解き方をさせることが可能になる。 ・複数の考え方を容認する実際のプログラミングに近い形での出題が可能になる。

プログラミング問題の例（PBT で出題する場合）

第●問 次のプログラムの空欄 **ア** ～ **オ** に入れる最も適当なものを、後の解答群のうちから一つずつ選べ。

トランプ 10 枚の数字（絵札の J は 11, Q は 12, K は 13 とする）が、次のように配列 Data に格納されている。この配列の値を昇順に並べ替えるバブルソート（隣接交換法）のプログラムを図 1 のように作成した。

なお、配列の要素を指定する添字は 0 から始まり、プログラムには次のような関数 **要素数** を用いている。

i	0	1	2	3	4	5	6	7	8	9
Data	11	7	5	1	5	12	7	6	3	7

要素数 (値) …配列の要素数を返す。

例：Data = [10, 20, 30, 40, 50, 60, 70] の時
要素数 (Data) は 7 を返す

(実行結果例)

1,3,5,5,6,7,7,11,12

```

(01) Data = [11, 7, 5, 1, 5, 12, 7, 6, 3, 7]
(02) saigo = 要素数(Data) - 1
(03) m を 0 から saigo - 1 まで 1 ずつ 増やしなが繰り返す:
(04) |   n を ア から イ まで 1 ずつ ウ なが繰り返す:
(05) |   |   もし エ ならば:      ア saigo          ア 0
(06) |   |   |   tmp = Data[n]      イ m+1          イ saigo-m-1
(07) |   |   |   Data[n] = オ      ウ 減らし        ウ 増やし
(08) |   |   |   オ = tmp
(09) 表示する (Data)      エ Data[n] < Data[n - 1]  エ Data[n] > Data[n + 1]
                        オ Data[n - 1]          オ Data[n + 1]

```

図1 バブルソートのプログラム

ア ～ **オ** の解答群

- | | | |
|-------------------------|-------------------------|-----|
| ① Data[n - 1] | ① Data[n + 1] | ② 0 |
| ③ 1 | ④ saigo | ⑤ m |
| ⑥ m+1 | ⑦ saigo-m-1 | |
| ⑧ 増やし | ⑨ 減らし | |
| Ⓐ Data[n] < Data[n - 1] | Ⓑ Data[n] > Data[n + 1] | |

プログラミング問題の例（CBT で出題する場合）

1 CBT イメージー 初期画面ー

The screenshot shows the initial interface of a CBT system. At the top, there's a header with '2020年度', '未解答数: 10 / 40', and a timer '残り時間: 1 時間 59 分 30 秒'. Below this, the problem title '設問 10 / 40問中' is displayed. The main content area is divided into three columns. The left column contains the problem description in Japanese, a table of data, and an example of the expected output. The middle column is a 'プログラミング' (Programming) area with a list of code snippets that can be dragged and dropped. The right column is a '解答 (実行結果)' (Answer (Execution Result)) area where the user can see the output of their code. A callout box with an arrow pointing to the programming area contains the text: '①短冊状のコードをプログラムエリアにドラッグ&ドロップしてプログラムを組み立てる' (Drag and drop the code snippets into the program area to build the program).

2020年度 未解答数: 10 / 40 設問一覧 操作方法 文字サイズ 残り時間: 1 時間 59 分 30 秒

設問 10 / 40問中

バブルソート

トランプ10枚の数字（絵札のJは11,Qは12,Kは13とする）が、次のように配列Dataに格納されている。この配列の値を昇順に並べ替えるバブルソート（隣接交換法）のプログラムを完成させなさい。

なお、配列の要素を指定する添字は0から始まり、プログラムには次のような関数要素数を用いている。

i	0	1	2	3	4	5	6	7	8	9
Data	11	7	5	1	5	12	7	6	3	7

要素数 (値) ...配列の要素数を返す。
例: Data=[10,20,30,40,50,60,70]の時
要素数(Data)は7を返す

(実行結果例)
1,3,5,5,6,7,7,11,12

プログラミング

右のプログラムエリアにドラッグ&ドロップする

```
Data[n] = tmp
Data[n - 1] = tmp
Data[n + 1] = tmp
tmp = Data[n]
Data[n] = Data[n - 1]
Data[n] = Data[n + 1]
n を saigo から m+1 まで 1 ずつ減らしながら繰り返す:
n を 0 から saigo-m-1 まで 1 ずつ増やしながら繰り返す:
もし Data[n] < Data[n - 1] ならば:
もし Data[n] > Data[n + 1] ならば:
```

Data = [11,7,5,1,5,12,7,6,3,7]

saigo = 要素数(Data)-1

m を 0 から saigo-1 まで 1 ずつ増やしながら繰り返す:

表示する(Data)

①短冊状のコードをプログラムエリアにドラッグ&ドロップしてプログラムを組み立てる

解答 (実行結果)

実行結果が解答となりますので、「実行」ボタンを押してここに結果を表示させてください。

実行

前へ 次へ

2 CBT イメージー プログラム完成・実行結果表示ー

This screenshot shows the same CBT interface as the previous one, but with the program completed and the execution results displayed. The programming area now shows the code snippets that were previously available, but they are no longer draggable. The right column shows the '解答 (実行結果)' (Answer (Execution Result)) with the output '1,3,5,5,6,7,7,11,12'. Callout boxes with arrows point to specific elements: '②短冊状のコードを組み立ててプログラムを完成させる。' (Build the program by combining the code snippets), '③実行ボタンをクリックする。' (Click the execution button), and '④実行結果を確認する。' (Check the execution result).

2020年度 未解答数: 10 / 40 設問一覧 操作方法 文字サイズ 残り時間: 1 時間 59 分 30 秒

設問 10 / 40問中

バブルソート

トランプ10枚の数字（絵札のJは11,Qは12,Kは13とする）が、次のように配列Dataに格納されている。この配列の値を昇順に並べ替えるバブルソート（隣接交換法）のプログラムを完成させなさい。

なお、配列の要素を指定する添字は0から始まり、プログラムには次のような関数要素数を用いている。

i	0	1	2	3	4	5	6	7	8	9
Data	11	7	5	1	5	12	7	6	3	7

要素数 (値) ...配列の要素数を返す。
例: Data=[10,20,30,40,50,60,70]の時
要素数(Data)は7を返す

(実行結果例)
1,3,5,5,6,7,7,11,12

プログラミング

右のプログラムエリアにドラッグ&ドロップする

```
Data[n] = tmp
Data[n - 1] = tmp
Data[n + 1] = tmp
tmp = Data[n]
Data[n] = Data[n - 1]
Data[n] = Data[n + 1]
n を saigo から m+1 まで 1 ずつ減らしながら繰り返す:
n を 0 から saigo-m-1 まで 1 ずつ増やしながら繰り返す:
もし Data[n] < Data[n - 1] ならば:
もし Data[n] > Data[n + 1] ならば:
```

Data = [11,7,5,1,5,12,7,6,3,7]

saigo = 要素数(Data)-1

m を 0 から saigo-1 まで 1 ずつ増やしながら繰り返す:

n を saigo から m+1 まで 1 ずつ減らしながら繰り返す: X

もし Data[n] < Data[n - 1] ならば: X

tmp = Data[n] X

n を 0 から saigo-m-1 まで 1 ずつ増やしながら繰り返す: X

もし Data[n] < Data[n - 1] ならば: X

もし Data[n] > Data[n + 1] ならば: X

②短冊状のコードを組み立ててプログラムを完成させる。

③実行ボタンをクリックする。

④実行結果を確認する。

解答 (実行結果)

1,3,5,5,6,7,7,11,12

実行

前へ 次へ

※ 上記のように、プログラムに関する選択肢があらかじめ与えられたプログラミング問題（CBT）であれば、示された実行結果例を得られれば正答、得られなければ誤答とするような採点基準を設定することもできるため、客観的で迅速な採点が求められる共通テストへの馴染みがよいと考えられる。

令和元年度実証実験（プレテストを想定）で取得した 解答データに基づく分析結果

○令和2年1～2月に、62問の問題から問題セット（8種類）を作成し、プレテストを想定した実証実験を実施（参加者は高校生307名、大学生340名）

○以下の基準の両方を満たした問題（太字・色付き）を問題バンクに登録。

- ・正答率（問題の難易度）：10%以上90%以下（太字）
- ・I-T相関係数*（問題の識別力）：0.20を超える（太字）

○各問に十分な参加者を割り当てることができなかったため、誤差（揺らぎ）が大きい値となっている。

問題番号	参加者数	正答率	I-T相関
11_1	161	.820	.381
11_2	161	.366	.198
11_3	161	.851	.131
11_4	160	.819	.154
12_1	163	.061	-.033
12_2	163	.951	.054
12_3	163	.503	.010
12_4	163	.994	.246
21_1	158	.595	.079
21_2	158	.228	.224
21_3	153	.752	.293
22_1	162	.765	.221
22_2	162	.062	.110
22_3	160	.425	.214
23_1	160	.806	.299
23_2	162	.889	.323
23_3	161	.578	.130
24_1	129	.752	.344
24_2	129	.264	.200
31_1	151	.411	.182
31_2	151	.907	.282
31_3	149	.687	.363
32_1	159	.648	.227
32_2	159	.730	.338
32_3	159	.585	.390
33_1	162	.747	.156
33_2	164	.951	.037
33_3	160	.562	.258
34_1	167	.275	.268
34_2	166	.458	.389
34_3	167	.820	.223

問題番号	参加者数	正答率	I-T相関
35_1	162	.735	.355
35_2	167	.602	.323
35_3	164	.613	.439
36_1	162	.284	.274
36_2	162	.562	.424
36_3	164	.415	.193
37_1	159	.528	.312
37_2	161	.950	.238
37_3	159	.579	.312
38_1	150	.587	.444
38_2	153	.876	.214
38_3	152	.921	.204
38_4	152	.921	.224
41_1	157	.146	.184
41_2	157	.382	.171
41_3	157	.420	.084
42_1	160	.306	.281
42_2	160	.400	.184
42_3	159	.937	-.098
43_1	165	.921	.208
43_2	165	.885	.208
43_3	163	.724	.192
44_1	166	.831	.114
44_2	166	.952	.103
44_3	165	.521	.355
45_1	165	.564	.206
45_2	165	.703	.373
45_3	161	.025	.198
46_1	156	.564	.264
46_2	156	.981	.305
46_3	155	.297	.265

※この分類は、あくまでも問題作成WG で作成した教科「情報」の問題に対して便宜上行ったものであり、妥当性や他教科への適応可能性などに
関してさらなる検討が必要である。

分類①（ $T_1 \cdot H_1$ ）種問題と派生問題の問題形式及び設定が同じであり、難易度もあまり変わらない。
曝露・漏洩の影響を考えると、派生問題の作成には慎重な検討が必要である。この問題例の場合、送受信の際に誤りが生じたビットの位置は異なるが、解答に当たっての考え方は同じであり、難易度も変わらない。

種問題（ T_1 ）例

問 パリティビットが付加された25ビットの受信データを、図1のように正方形に並べた。このとき、行と列それぞれに奇数パリティになるようなパリティビットが付加されて送られていた場合、送受信の際に誤りが生じたビットの位置の列と行を選べ。

	A	B	C	D	E
①	1	0	1	1	0
②	0	1	1	0	1
③	1	1	1	1	0
④	0	0	1	0	0
⑤	1	1	1	0	0

図1 パリティビットが付加された受信データ

派生問題（ H_1 ）例

問 パリティビットが付加された25ビットの受信データを、図1のように正方形に並べた。このとき、行と列それぞれに奇数パリティになるようなパリティビットが付加されて送られていた場合、送受信の際に誤りが生じたビットの位置の列と行を選べ。

	A	B	C	D	E
①	1	0	1	1	0
②	0	1	1	0	1
③	1	1	1	0	0
④	0	0	1	0	1
⑤	1	1	1	0	0

図1 パリティビットが付加された受信データ

分類②（ $T_2 \cdot H_2$ ）種問題と派生問題の問題形式及び設定が同じであるが、与える条件によって難易度を変えることができる。
曝露・漏洩の影響を考えると、派生問題の作成には慎重な検討が必要である。この問題例の場合、種問題では解答時に結合すべきテーブルが二つであるのに対して派生問題では三つであり、派生問題の方がやや難易度が高くなる。

種問題（ T_2 ）例

問 コンビニエンスストアで、ポイントカード会員を、次に示すテーブルからなるデータベースを用いて管理している。次の操作を行ったときに、現在のテーブルとして得られるテーブルを①～④のうちから一つ選べ。

対象 商品テーブル
選択 分類 ジュース
結合 購入履歴テーブル
射影 商品名 単価 個数
(操作の説明は省略)

会員テーブル

会員番号	名前	電話番号	生年月日	住所
20200001	駒場 大	03-34**-****	200x.4.2	東京都目黒区
20200002	目黒 学	03-34**-****	200x.3.31	東京都目黒区
20200003	文科 省一	03-52**-****	200x.12.31	東京都千代田区

商品テーブル

商品番号	分類	商品名	単価	付与ポイント
J10000001	ジュース	オレンジジュース500mL	140	14
J10000002	ジュース	リンゴジュース1L	380	38
T20000001	お茶	緑茶500mL	130	13

購入履歴テーブル

会員番号	商品番号	個数	販売日時
20200002	J10000002	1	20xx/02/23 09:13:27
20200001	J10000001	2	20xx/02/23 09:21:41
20200003	T20000001	1	20xx/02/23 09:35:16

【解答群】

①	名前	商品名	単価	個数
	目黒 学	リンゴジュース1L	380	1
	駒場 大	オレンジジュース500mL	140	2

派生問題（ H_2 ）例

問 コンビニエンスストアで、ポイントカード会員を、次に示すテーブルからなるデータベースを用いて管理している。次の操作を行ったときに、現在のテーブルとして得られるテーブルを①～④のうちから一つ選べ。

対象 購入履歴テーブル
結合 会員テーブル
選択 住所 東京都目黒区
結合 商品テーブル
射影 名前 商品名 単価 個数
(操作の説明は省略)

会員テーブル

会員番号	名前	電話番号	生年月日	住所
20200001	駒場 大	03-34**-****	200x.4.2	東京都目黒区
20200002	目黒 学	03-34**-****	200x.3.31	東京都目黒区
20200003	文科 省一	03-52**-****	200x.12.31	東京都千代田区

商品テーブル

商品番号	分類	商品名	単価	付与ポイント
J10000001	ジュース	オレンジジュース500mL	140	14
J10000002	ジュース	リンゴジュース1L	380	38
T20000001	お茶	緑茶500mL	130	13

購入履歴テーブル

会員番号	商品番号	個数	販売日時
20200002	J10000002	1	20xx/02/23 09:13:27
20200001	J10000001	2	20xx/02/23 09:21:41
20200003	T20000001	1	20xx/02/23 09:35:16

【解答群】

①	名前	商品名	単価	個数
	目黒 学	リンゴジュース1L	380	1
	駒場 大	オレンジジュース500mL	140	2

分類③ (T₃・H₃) 種問題と派生問題の問題形式は同じであるが、設定及び条件を変えることで難易度を
変えることができる。

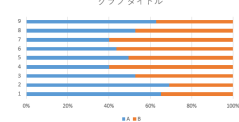
問題が曝露された場合でも設定や条件が異なればある程度派生問題を作成することができる。この例では、グ
ラフにすることを考えることによって複数の派生問題を作ることができ、また難易度も変えることができる。

種問題 (T₃) 例

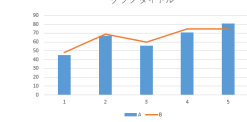
問 Aさんは、次の内容を表すグラフを検討している。最も適切なグラフを次の①～⑧のうちから一つ選べ。ただし、図表はイメージ図であり、実際の数値等は表されていない。

「『スマホの使用時間が多い人は、勉強時間が少ないのではないか』ということを調べるために、スマホの使用時間と勉強時間の関係性を表したい。」

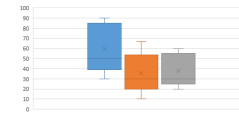
①100%積み上げ横棒グラフ



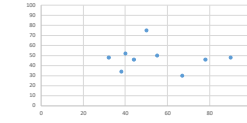
②複合グラフ (棒グラフ, 折れ線グラフ)



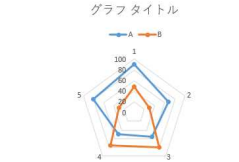
③箱ひげ図



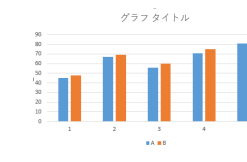
④散布図



⑤レーダーチャート



⑥縦棒グラフ



派生問題 (H₃) 例

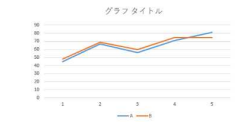
問 Aさんは、次の内容を表すグラフを検討している。最も適切なグラフを次の①～⑧のうちから一つ選べ。ただし、図表はイ
メージ図であり、実際の数値等は表されていない。

「医師の地域的な偏りを調べるために、各都道府県の人口10万人あたりの医師数 (医療施設に従事する医師数) について、都道府県別に過去10年間のばらつきを表したい。」

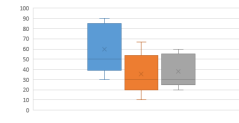
①円グラフ



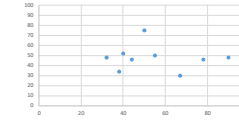
②折れ線グラフ



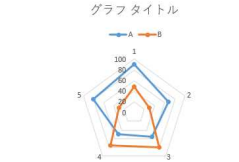
③箱ひげ図



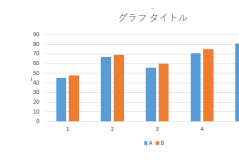
④散布図



⑤レーダーチャート



⑥縦棒グラフ



「情報」試作問題（検討用イメージ）（抜粋）

大学入試センターは、新しい学習指導要領に対応した大学入学共通テストの新しい科目として「情報」を検討しており、大学や高等学校等の関係者に御意見を伺いながら検討を進めている。その中で、大学入試センターでは、共通テスト「情報」の試験問題について具体的なイメージを共有するために、「情報」試作問題（検討用イメージ）を大学や高等学校等の関係者に提供しているが、その試作問題の一部（以降の4問）は、CBT 問題作成ワーキンググループにおいて CBT の問題として作成しものを改変したものである。

本件は「情報」の出題が決まったものではありませんが、高校・大学関係団体が大学入学共通テストの出題科目のあり方について検討できるよう、その参考として提供するものです。

「情報」試作問題（検討用イメージ）

本冊子の趣旨

※本冊子をご覧になる前に必ずお読みください※

- この冊子の試作問題群（以下「本試作問題群」という。）は、大学入学共通テストへの導入を検討している「情報」について具体的なイメージを共有するために、(独)大学入試センターにて用意したものです。今後、大学や高等学校等の関係者に御意見を伺いながら、大学入学者選抜としての適切な出題について引き続き検討することとしています。
- 本試作問題群は、平成 30 年に改訂された高等学校学習指導要領（「情報Ⅰ」）に基づいて作成したものです。「情報Ⅰ」のできる限り多くの項目を網羅できるように、また様々な問題形式の可能性を提示するために、多様な試作問題を掲載しています。「情報Ⅰ」については、次ページ以降の解説も御覧ください。
- 本試作問題群は、検討用イメージとして作成したものであるため、活用にあたっては以下の点に十分御留意いただきますようお願いします。
 - ※ 多様な試作問題を掲載していますが、「情報Ⅰ」の全ての項目を網羅しているものではありません。
 - ※ 「情報Ⅰ」の教科書は現在検定中ですので、本試作問題の内容は教科書と照合したものではありません。
 - ※ 本試作問題は専門家による検討を経たものですが、過去のセンター試験や大学入学共通テストと同様の問題作成や点検のプロセスを経たものではなく、また、実際の問題セットをイメージしたものや試験時間を考慮したものでもありません。仮に「情報」が出題科目となる場合には、適切な分量と難易度のもとで問題セットが作成されることとなります。
 - ※ 新たに作成した問題がほとんどですが、一部に、過去のセンター試験の「情報関係基礎」で出題した問題の改題を含んでいます。
- 10月20日付け入試セ企第74号の別添「平成30年告示高等学校学習指導要領に対応した大学入学共通テストの出題教科・科目について（検討中案）」において、「令和7年度大学入学共通テストではPBT（Paper-based Testing：紙で実施する試験）で行うことを基本としつつ、現在進めているCBT（Computer-based Testing:コンピュータ等で実施する試験）に関する調査研究の状況を踏まえ検討する」とされています。

解説

情報のデジタル化に関する仕組みや情報量に関する考え方は、情報技術を活用する上で基本となる知識・技能である。

この問題は、情報の表し方や身近な動画のデータ量に関する基本的な知識・技能を問うものである。

対応する情報Ⅰの主な領域：(2) コミュニケーションと情報デザイン 問題種：小問

第2問 後の各問いに答えよ。

問2

Mさんはスマートフォンで動画を撮りたいと考えた。しかし、スマートフォンのデータ保存用メモリの空き容量が足りるか心配になったため、動画撮影アプリの設定画面（図1）で画像サイズ等の設定を変えることでデータ量を小さくしたいと考えた。

次のⅠ～Ⅲの設定で撮影された1秒あたりの動画のファイルサイズを、小さい順に不等号で区切り並べたものを次の①～⑤のうちから一つ選べ。なお、圧縮などは考えないものとする。

ア

ア ①Ⅰ<Ⅱ<Ⅲ



図1 動画の設定画面

表1 動画撮影の設定

記号	色数	フレームレート	画像サイズ(ピクセル)
Ⅰ	16,777,216 色(24bit)	60fps	1280×720
Ⅱ	16,777,216 色(24bit)	30fps	1920×1080
Ⅲ	256 色	30fps	3840×2160

ア

の解答群

- ① Ⅰ<Ⅱ<Ⅲ ② Ⅰ<Ⅲ<Ⅱ ③ Ⅱ<Ⅰ<Ⅲ ④ Ⅱ<Ⅲ<Ⅰ
⑤ Ⅲ<Ⅰ<Ⅱ ⑥ Ⅲ<Ⅱ<Ⅰ

解説

学んだ情報技術がどのように社会の中で利用されているかを結び付けて考える必要がある。これは、菓子メーカーの工場における不良品の判別に利用される画像処理（画像を白と黒の2色のみで表現する2階調、明度と画素数等）に関する問題である。

対応する情報Ⅰの主な領域：(2)コミュニケーションと情報デザイン 問題種：中問

第3問 次の文章を読み、空欄 **ア** ～ **ウ** に入れるのに最も適当なものを、文の後の解答群のうちから一つ選べ。

ある菓子メーカーの工場では、出来上がったせんべいを袋詰めする前に製造ライン上でカメラ撮影して、割れや欠けなどの不良品の検出を自動で判別する装置を導入している。装置は、割れや欠けがあるせんべいを判別しやすいように撮影した画像を白と黒の2階調に変換（二値化）して処理をしている。図1は写真1の画素を明度でヒストグラムに表したものである。二値化を行う際の濃度変換の分かれ目となる濃度値（しきい値）をAとBとした場合、しきい値Aの時の画像は **ア** であり、しきい値Bの時の画像は **イ** となる。

また、この装置では割れや欠けがあるせんべいを判別しやすいようにプログラムで自動的に二値化のしきい値を決めている。図1のように、明度と画素数のヒストグラムにおいて二つの山型があった場合、最適なしきい値は **ウ** と判断することができる。

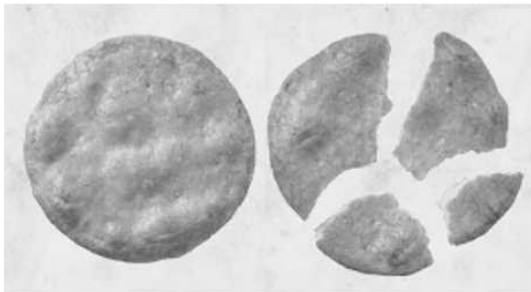


写真1 せんべいの写真

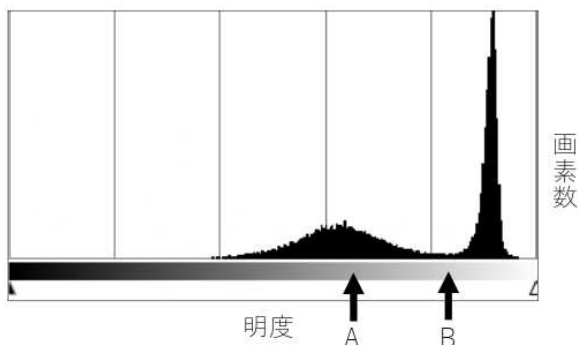


図1 明度と画素数

ア・イの解答群

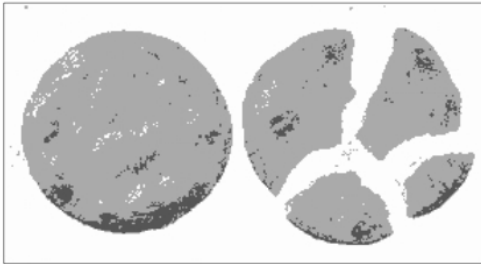
①



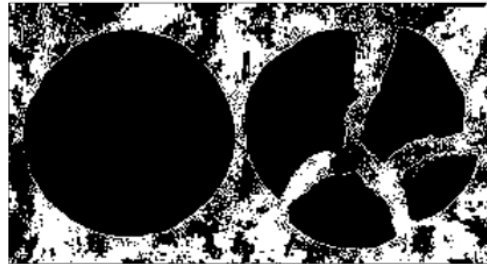
②



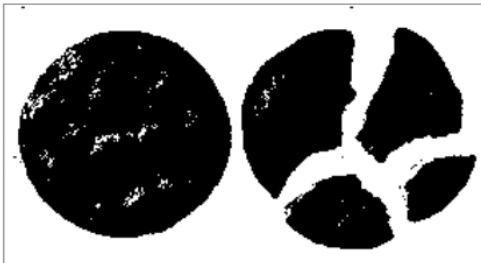
③



④



⑤



⑥



ウの解答群

- ① 低い山の最も画素数が多い明度
- ② 高い山の最も画素数が多い明度
- ③ ヒストグラムの中央値の明度
- ④ ヒストグラムの平均値の明度
- ⑤ 二つの山の間の最も画素数が少ない明度
- ⑥ 山に関係なく画素数が最も多い明度

解説

デジタル社会においては、情報セキュリティを正しく理解していることは生きていく上で必須である。

この問題は、社会でも一般的に使われるようになった二要素認証の情報セキュリティ上の有用性に関する正しい理解を問うものである。

※二要素認証は、報道等で頻繁に取り上げられた「二段階認証」の一種である。この問題では、敢えて「二要素認証」という、より正確な用語を導入・解説している。

対応する情報Ⅰの主な領域：(4)情報通信ネットワークとデータの活用 問題種：小問

第 6 問

クラウド上の決済サービスなどでは、より強固な認証が必要である。そのため、近年はスマートフォンを利用した二段階認証の一種である二要素認証が使われる場合も多い。これは例えば、これまでの ID とパスワードに加えて、利用者本人が事前に登録したスマートフォンに送信される一度限り有効なパスワードを用いる方法である。次の図中の 1～4 は、この二要素認証の手順を模式的に表したものである。

この二要素認証によって、セキュリティが強固になる理由として最も適切なものを次の ①～④のうちから一つ選べ。 ア ア ②

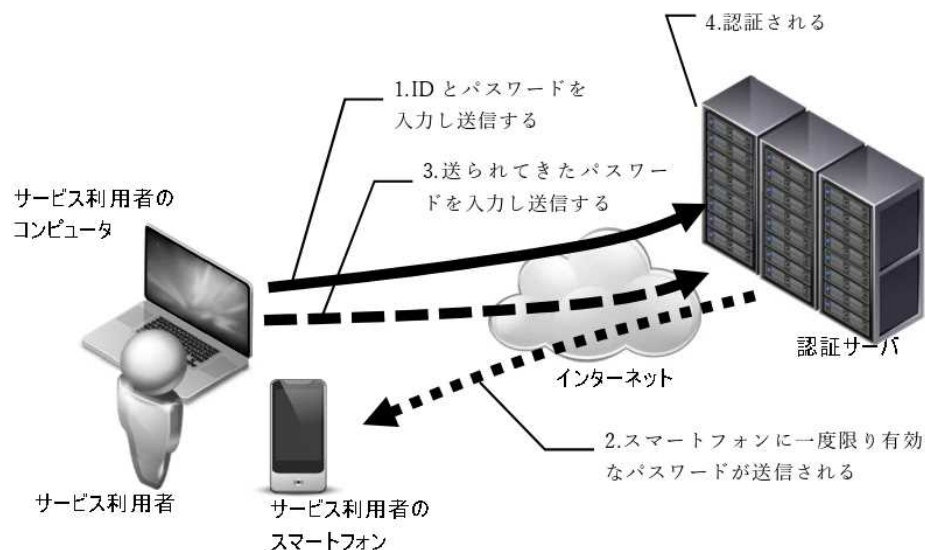


図 二要素認証の模式図

アの解答群

- ① 利用するサイトが正しいサイトであれば、入力したパスワードがスマートフォンに送信されるため
- ② パスワードを2回、時間をあけて入力して認証するため
- ③ IDとパスワードを知っていることに加え、登録されたスマートフォンを持っていることを確認できるため
- ④ IDとパスワードを知っていることに加え、スマートフォンのGPS機能を使って居場所を特定して認証するため

解説

家庭や学校にも情報通信ネットワークは普及しており，その仕組みを正しく理解し，トラブルなどに対応できる力が必要とされている。この問題は，通信状況からネットワークの不具合の原因を推定する力を問うものである。

※新学習指導要領では，情報通信ネットワークの具体的な指導内容として，「実際に家庭内 LAN 等の小規模な情報通信ネットワークを構築したり，あらかじめ用意したトラブルを抱えている情報通信ネットワークの不具合を解決したりすることを扱うことも考えられる。」とあり，その内容を踏まえた問題。

対応する情報Ⅰの主な領域：(4)情報通信ネットワークとデータの活用 問題種：中間

第7問 次の文章を読み，空欄に入れるのに最も適当なものを後の解答群のうちから一つ選べ。

高校生の T さんは，放課後に調べものをするため，視聴覚室にあるパソコンでインターネットに接続しようとしたところできなかった。T さんの高校におけるネットワークの構成は，次の図 1 のようになっている。

T さんはコンピュータなどの管理を手伝っていたので早速不具合の原因を調べることにした。まず，視聴覚室のパソコンからいくつかの IP アドレスにパケットが届くかを確認（疎通確認）したところ，表 1 のようになった。スイッチングハブまたはルータのいずれかが 1 台故障したと考えると，故障の可能性がある機器は **ア** と **イ** である。 **ア ②C のスイッチングハブ**
イ ①A のスイッチングハブ

次に，このどちらが故障しているかを判別するために，1 年 1 組の教室に移動して教室内にあるアクセスポイントに接続したタブレット端末から疎通確認を行った。ここで，**ウ** にパケットが届けば **ア** が故障と特定でき，パケットが届かなければ **イ** が故障と特定できる。 **ウ ③192.168.1.11**

表 1 視聴覚室からの疎通結果

送信先	結果
192.168.1.1	×
192.168.1.11	×
192.168.1.21	×
192.168.1.31	×
192.168.1.61	○
192.168.1.101	○

○：パケットが届く ×：パケットが届かない

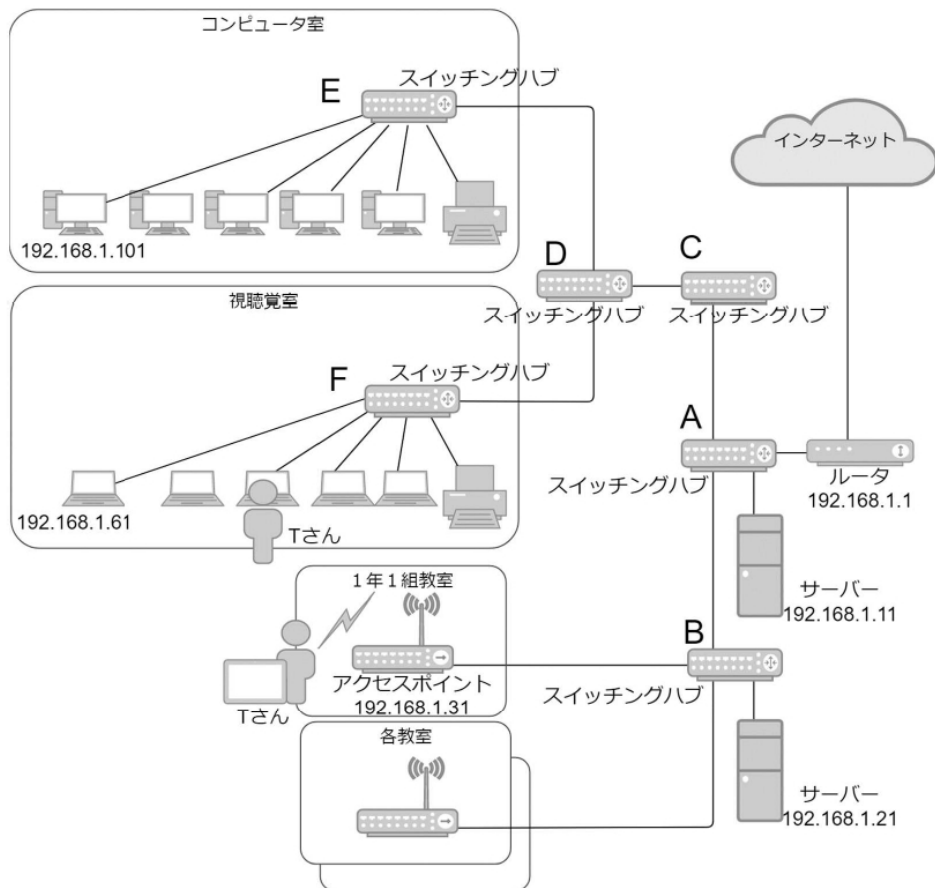


図1 ネットワークの構成図

ア・イの解答群

- | | |
|---------------|----------------|
| ① A のスイッチングハブ | ① B のスイッチングハブ |
| ② C のスイッチングハブ | ③ D のスイッチングハブ |
| ④ E のスイッチングハブ | ⑤ F のスイッチングハブ |
| ⑥ ルータ | ⑦ この情報では特定できない |

ウの解答群

- | | | |
|----------------|-----------------|----------------|
| ① 192.168.1.11 | ① 192.168.1.21 | ② 192.168.1.31 |
| ③ 192.168.1.61 | ④ 192.168.1.101 | |

大学入学共通テスト企画委員会 CBT 活用検討部会委員（五十音順）

（職名は令和 3 年 3 月現在）

◎：部会長，○：副部会長

植 野 真 臣	電気通信大学大学院情報理工学研究科教授
宇佐美 慧	東京大学高大接続研究開発センター准教授
◎ 川 嶋 太津夫	大阪大学高等教育・入試研究開発センター長・特任教授
澤 木 泰 代	早稲田大学教育・総合科学学術院教授
白 水 始	国立教育政策研究所初等中等教育研究部総括研究官
千 代 勝 実	山形大学学士課程基盤教育機構教授
高 井 潤	埼玉県立狭山工業高等学校主幹教諭
立 脇 洋 介	九州大学アドミSSIONセンター准教授
椿 美智子	電気通信大学大学院情報理工学研究科教授
西 郡 大	佐賀大学アドミSSIONセンター教授
仁 田 善 雄	公益財団法人 医療系大学間共用試験実施評価機構理事
萩 原 兼 一	大阪大学名誉教授，同大学大学院情報科学研究科招へい教授
萩 谷 昌 己	東京大学大学院情報理工学系研究科教授
林 篤 裕	名古屋工業大学大学院工学研究科教授
○ 前 川 眞 一	（独）大学入試センター特任教授，東京工業大学名誉教授
室 田 真 男	東京工業大学リベラルアーツ研究教育院教授

（オブザーバー）

浅 原 寛 子	文部科学省総合教育政策局調査企画課学力調査室長
佐 藤 安 紀	国立教育政策研究所次長
前 田 幸 宣	文部科学省高等教育局大学振興課大学入試室長

その他の協力者・団体（五十音順）

大 塚 尚 子 国立教育政策研究所国際研究・協力部総括研究官
大学入学共通テスト企画委員会 CBT 活用検討部会 CBT 問題作成ワーキンググループ（※）
日本タタ株式会社
プロメトリック株式会社

（※）委員名は非公表

(独)大学入試センター

白 井 俊	(独)大学入試センター試験企画部長
平 千 枝	(独)大学入試センター試験企画部試験企画課参事
上遠野 奈保子	(独)大学入試センター試験企画部試験企画課長
西 野 努	(独)大学入試センター試験企画部試験企画課課長補佐
甲 斐 一 成	(独)大学入試センター試験企画部試験企画課試験企画第三係長
田 畑 さくら	(独)大学入試センター試験企画部試験企画課企画調整係長
中 尾 ちひろ	(独)大学入試センター試験企画部試験企画課試験企画第三係
垣 内 将 平	(独)大学入試センター試験企画部試験企画課企画調整係
水 野 修 治	(独)大学入試センター試験問題調査官 (情報担当)
桑 山 一 也	(独)大学入試センター試験問題調査官 (配慮担当)
前 川 眞 一	(独)大学入試センター特任教授, 東京工業大学名誉教授 (再掲)
寺 尾 尚 大	(独)大学入試センター研究開発部助教
荘 島 宏二郎	(独)大学入試センター研究開発部准教授
南 谷 和 範	(独)大学入試センター研究開発部准教授
宮 澤 芳 光	(独)大学入試センター研究開発部助教
大 塚 雄 作	(独)大学入試センター名誉教授/客員教授, 国際医療福祉大学大学院特任教授