

全国的な学力調査の CBT 化検討ワーキンググループ

最終まとめ

令和3年7月16日

全国的な学力調査の CBT 化検討ワーキンググループ

全国的な学力調査の CBT 化検討ワーキンググループ

最終まとめ

令和3年7月16日

<総論>

- 現在、文部科学省では、学校における1人1台端末と高速大容量の通信ネットワークを一体的に整備し、それらを活用して全ての子供たちの可能性を引き出す、個別最適な学びと協働的な学びの実現に向けた GIGA スクール構想¹を推進している。「全国的な学力調査の CBT²化検討ワーキンググループ」(以下「本 WG」という。)は、この GIGA スクール構想を踏まえ、全国学力・学習状況調査を CBT として実施する構想について、専門的・技術的観点から検討を行うことをその役割として、令和2年4月に「全国的な学力調査に関する専門家会議」(以下「専門家会議」という。)の下に設置され、同年5月から本格的に検討を開始し、全11回にわたって、有識者等からのヒアリングを行いつつ、検討を行ってきた。
- 国際学力調査である PISA³ (OECD 生徒の学習到達度調査) は、2015年調査より従来の PBT⁴から CBT に全面移行した。TIMSS⁵ (国際数学・理科教育動向調査) においても、2019年調査より、従来の PBT に加え CBT が一部導入され、2023年調査では全面移行することが予定されている。CBT による学力調査が国際的な標準となりつつある中で、我が国においても、CBT の特性・利点を活かした出題等、調査の一層の質の向上を図るとともに、教育データの収集・分析・利活用の充実による EBPM⁶ のさらなる推進を図るため、国際的に後れをとることなく、全国学力・学習状況調査の CBT 化に向けた検討・取組を進めることが急務である。
- 本 WG の中間まとめ(令和2年8月28日)では、CBT 化の検討をさらに進めていくにあたり、「調査の目的とその実施方法は表裏一体の関係である」として、調査の目的や役割等について総合的に検討する必要性について指摘した。
- この指摘を受けて、専門家会議において、全国学力・学習状況調査のこれまでの成果と課題、また、独自に調査を実施している地方自治体からのヒアリングなどを踏まえ議論が行われた。その上で、専

¹ GIGA スクール構想：Society 5.0 時代を生きる子供たちにとって、教育における ICT を基盤とした先端技術等の効果的な活用が求められる中、我が国の遅れた学校 ICT 環境を抜本的に改善すべく、ハード面の児童生徒1人1台端末や高速大容量の通信ネットワークなどの整備とともに、ソフト面ではデジタル教科書・教材等のデジタルコンテンツの活用促進、人材面では GIGA スクールサポーターの配置等、一体的な取組を推進。令和3年3月末時点で、全自治体等のうち1,748自治体等(96.5%)が令和2年度内に端末の納品を完了する見込み。

² CBT (Computer Based Testing)：コンピュータ使用型調査

³ PISA (Programme for International Student Assessment)：OECD (経済協力開発機構)による国際的な学習到達度に関する調査。日本では高等学校1年生等が対象。

⁴ PBT (Paper Based Testing)：筆記型調査

⁵ TIMSS (Trends in International Mathematics and Science Study)：IEA (国際教育到達度評価学会)による国際共同研究調査。日本では小学校4年生、中学校2年生が対象。

⁶ EBPM (Evidence-Based Policy Making)：エビデンスに基づく政策立案

専門家会議として、①毎年、原則として悉皆で実施している本体調査（以下、「悉皆調査」という。）と、②それを補完する調査として、IRT（項目反応理論）⁷を採用し、抽出により全国的な学力の状況について経年の変化を把握する経年変化分析調査⁸及び保護者に対する調査⁹（以下、「経年調査」という。）を、いわば国が実施すべき主要な調査の「二本柱」として位置づけ直し、整理することが提言された¹⁰。

○ 具体的には、「地方自治体や学校に対して、学習指導要領の理念や目標、内容等を具体的に示し、児童生徒の学習指導の改善や地方の教育施策の検証に資すること」を目的とした悉皆調査と、「国レベルで時系列的な学力の変化を正確に観察すること」を目的とした経年調査とで、適切に役割分担を図りつつ、それぞれの目的に即して最適な方法を設計し、CBT化していくことが必要であるとの方向性が示された¹¹。

○ 本WGでは、専門家会議から示されたこの方向性を踏まえ、国が実施すべき学力調査、すなわち、悉皆調査及び経年調査のCBT化について、それぞれの調査の目的に即して、専門的・技術的な観点から論点整理を行い、最終まとめとする。本WGにおいて検討した諸課題に関する具体的な内容は、調査の目的に即して「各論」として論点ごとにまとめているが、各論点における指摘事項は、相互に他の論点における事項にも影響を及ぼす面があることに留意されたい。また、全体を通じて、特に留意すべき点については、以下（1）～（5）のとおりである。

なお、本最終まとめにおける用語の整理として、「全国学力・学習状況調査」とは、悉皆調査（児童生徒質問紙調査を含む。）及び経年調査を総称するものとし、各調査固有の内容について指摘する場合は、それぞれの調査名を明示している。

（1）段階的な試行・検証の必要性

全国学力・学習状況調査のCBT化にあたっては、すでにCBTで実施されている学力調査や試験の事例等も参考に、まずは小規模でCBTの試行とその検証に取り組み、その結果や明らかになった課題を、次回の試行・検証に向けた改善に生かすというPDCAサイクルにより前進させていく取組が現実的かつ合理的である。文部科学省においては、まず、令和3年度から約100校を対象に試行・検証

⁷ IRT（Item Response Theory）：問題の特性（難易度や識別力等）と各児童生徒の学力を分けて考える枠組みであり、例えば、児童生徒の正答・誤答が、問題の難易度（簡単・難しい）によるのか、問題の質によるのか（測りたい能力を正確に測れる問題か）、児童生徒の学力によるのかなどを区別して考えることができる。これにより、いくつかの手続きを経ることで、異なる時点、冊子、児童生徒集団等で実施した場合であっても、その結果を相互に比較できるようになる。PISAやTIMSS、医療系大学間共用試験等で採用されている。

⁸ 経年変化分析調査：全国的な学力の状況について、経年の変化をテスト理論に基づいて把握・分析し、今後の教育施策の検証・改善に役立てることを目的とした調査。これまでに3回（平成25年度、平成28年度、令和3年度）実施している。令和3年度は、統計的手法に基づき厳格な抽出を行った結果選ばれた国・公・私立の小学校第6学年、中学校第3学年の児童生徒（全国で小学校600校程度、中学校750校程度）を対象に実施。

⁹ 保護者に対する調査：家庭状況と学力等の関係について、経年の変化を把握・分析し、今後の教育施策の検証・改善に役立てることを目的とした調査。これまでに3回（平成25年度、平成29年度、令和3年度）実施している。

¹⁰ 令和3年3月22日開催「第8回全国的な学力調査に関する専門家会議」耳塚座長ご発表資料

¹¹ あわせて、国際学力調査に引き続き参加する必要性や、地方自治体が独自に実施する個々の児童生徒の学力の発達等を継続的に測定して指導改善に活用することを目的とした調査について、国として必要な支援を検討することの必要性も指摘されている。

を開始することとしている¹²。全国レベルでの実施に向けては、ネットワーク・システムや学校での実施体制及び学校において必要となるサポート体制、問題表示形式等による差異、CBTとPBTとの解答（回答）の違い、合理的配慮の在り方等、様々な観点から検証を行い、課題の抽出とその解決を繰り返しつつ、段階的にその規模・内容を拡充しながら、全国レベルでの実施に向けて取り組んでいくことが必要である。

（２）端末による学習環境への習熟と発達段階等への考慮

児童生徒の学力を的確に測るためには、児童生徒が端末の基本的な操作や端末を用いた学習に習熟した上で、端末を用いた調査に違和感なく取り組めるようになることが重要である。令和3年度以降、GIGAスクール構想による様々な取組によって、学校における日常的なICTを活用した学習が本格的に推進されていくことも踏まえ、児童生徒が普段の学習で使い慣れている環境で調査を実施できるようにしていくことが必要である。また、例えば小学生が記述式問題に解答する際のキーボードの使用や障害のある児童生徒への配慮や指導について、十分な考慮が必要である。

（３）CBT化による学校現場への負担の考慮

全国学力・学習状況調査のCBT化にあたっては、地方自治体及び学校等によってネットワーク構成等が異なることを踏まえ、学校の端末等から国のCBTシステムまでを円滑に接続できる環境の整備・検証や、調査実施前における各端末の動作確認や不具合等の現場での対応等、地方自治体及び学校等への新たな負担をできる限り軽くすることなどについて、十分な配慮が必要である。具体的には、各学校や教職員に対し技術的な支援を行うため、ICT活用教育アドバイザーやICT支援員等の配置を進めるとともに、必要な人材養成を進める必要がある。さらに、CBTによる調査の実施にあたって、技術面から各学校等をサポートする体制を構築することも必要である。

（４）詳細な調査設計の検討の必要性

各調査の目的を踏まえ、悉皆調査については、教育指導や地方自治体における施策改善への活用等の観点から、また、経年調査については、全国的な学力の状況の経年の変化をより効率的に測定する観点から、問題作成及び結果分析の枠組みの在り方等をはじめ、学習科学や教育測定、教科教育学、データ科学等の幅広い知見を有効に生かし、詳細に検討する必要がある。本最終まとめを基に、試行・検証を通じて判明した課題等も踏まえつつ、目的に即した的確な調査設計となるよう、各分野の専門家等とともに検討しながら、改善を重ねていくことが必要である。

¹² 令和3年度の試行・検証の規模は小学校第6学年、中学校第3学年の児童生徒約1万人。主な検証事項は、（１）ネットワーク・システムの検証（①ネットワーク環境、②ネットワークの負荷、③CBTシステムのサーバの負荷）、（２）学校での実施体制やサポート体制の検証、（３）問題の検証（①問題表示形式等による差異、②児童生徒の文字入力能力）等。

(5) CBT 化に向けた体制整備の必要性

全国学力・学習状況調査の CBT 化に向けて、文部科学省及び国立教育政策研究所がシステムや作問体制等の検討を含め先導していく必要がある。こうした大規模調査の CBT 化にあたっては、前述のとおりきめ細かな配慮をしながら、調査の目的等に即した詳細な調査設計の検討や必要なシステム機能に関する研究、十分な質・量の調査問題の作成等、膨大な作業を行っていく必要がある。また、教育 DX¹³の動向に対応し、CBT のメリットを最大化していくためには、あわせてテスト理論やデータ分析に精通している人材を確保し、積極的に活用する必要がある。

なお、国立教育政策研究所においては、CBT 化に向けた検討・準備の一環として、令和3年10月に「教育データサイエンスセンター」が新設される予定であるが、今後、CBT への移行及び安定的な事業運営のため、同研究所の体制をさらに強化していくことが不可欠である。その際、大学や教育委員会等との連携体制の構築についても進めていくことが必要である。

- 全国学力・学習状況調査の CBT 化の実現のためには、地方自治体及び学校のネットワーク環境の整備等、実施に向けた準備が必要となる。本最終まとめでは、本 WG 及び専門家会議での議論や文部科学省及び国立教育政策研究所における取組等を踏まえ、地方自治体や学校が CBT 化に向けた準備を計画的かつ段階的に進めることができるよう、「工程に関する考え方」を示すこととする。文部科学省及び国立教育政策研究所においては、この考え方を踏まえつつ、地方自治体及び学校等の理解と協力を丁寧に得ながら、GIGA スクール構想の進捗状況も踏まえつつ、CBT 化に向けた取り組みを着実に進めていただきたい。

¹³ DX (Digital Transformation) : 将来の成長、競争力強化のために、新たなデジタル技術等を活用して新たなビジネスモデルを創出・柔軟に改変すること。

<各論>

(1) 日々の学習における ICT の活用と全国的な学力調査との関係

(日々の学習における ICT の活用の必要性)

- GIGA スクール構想を踏まえ、全国規模の学力調査を CBT で行うためには、端末や安定したネットワークが全国的に整備されることが大前提である。その上で、測りたい学力を的確に測定するために児童生徒や教師が日常的に ICT 機器を使用した学習やパソコン操作に慣れていることが重要である。
- PISA2018 において、我が国の児童生徒は、学習の場面で ICT を活用した学習活動を学校内外で十分に行っていないことが判明しており、ICT 環境の整備や活用の遅れが課題となっている¹⁴。実際、TIMSS2019 において、試行的に CBT で調査を実施した際、一部の生徒がローマ字入力や入力モードの切り替え（数字や元素記号等）でつまづく様子がみられた。また、平成 31（令和元）年度に実施した全国学力・学習状況調査の中学校英語「話すこと」調査でも、生徒がコンピュータに向かって口述で解答する調査方法に慣れておらず、操作に戸惑い、調査開始が遅れるなどのケースがあった。
- 新学習指導要領では、情報活用能力¹⁵について、言語能力等と同様に、「学習の基盤となる資質・能力」と位置づけられており、教育課程全体で確実に育むこととされている。また、GIGA スクール構想により、1人1台端末が配備されるなど、学校現場における ICT 環境の整備が加速度的に進められており、今後、児童生徒が学校内外で端末を活用しながら使い慣れていくことが想定される。さらに、令和3年度から学びの保障オンライン学習システム（MEXCBT：メクビット）¹⁶の開発が進められており、児童生徒が端末上で様々な形式の問題に解答する形に慣れていくことも期待される。
- 全国学力・学習状況調査の CBT 化にあたっては、測りたい学力と児童生徒の端末の基本的な操作等の関係について、令和3年度から実施する試行・検証や情報活用能力調査の結果等も踏まえつつ、把握・分析していくことが必要である。令和3年度から文部科学省が実施する CBT 化に向けた試行・検証においては、まず児童生徒のキーボードなど¹⁷による文字入力数¹⁸を測定した上で、短答式の解答

¹⁴ PISA2018 の結果によると、日本は、OECD 加盟国中で学校の授業（国語、数学、理科）におけるデジタル機器の利用時間が最も短い。

¹⁵ 情報活用能力：世の中の様々な事象を情報とその結び付きとして捉え、情報及び情報技術を適切かつ効果的に活用して、問題を発見・解決したり自分の考えを形成したりしていくために必要な資質・能力。

¹⁶ MEXCBT（メクビット）：文部科学省が構築している、児童生徒が学習端末を用いてオンラインで問題演習等ができる CBT システム（問題やデータの相互運用が可能な国際標準規格に基づく汎用的なシステム）。令和2年度はプロトタイプを開発。令和3年度から、希望する全国の小・中・高校等で活用可能となっている。

¹⁷ 学校における ICT 環境整備の在り方に関する有識者会議の「最終まとめ」（平成29年8月）において、「これからの学習活動を支える ICT 機器等の機能」として、学習者用コンピュータ（児童生徒用）には、「キーボード「機能」を有することが必要。※小学校中学年以上では、キーボードを必須とすることが必要」と指摘されている。

¹⁸ 情報活用能力調査：平成25年度の結果によると、キーボードによる文字入力数は、小学校（第5学年）は5.9文字/分（参考値）、中学校（中学校第2学年）は15.6文字/分。

結果とあわせて分析することとしている。以降も、試行の内容を充実させながら、学力と児童生徒の端末の操作等の関係について、継続的に把握・分析していく必要がある。

(2) CBTの利点を活かした学力調査の在り方

① 調査日程・期間

(一定期間内(複数日に分散)の実施)

- 現行の悉皆調査は、全国同日一斉実施で行われており、実施日までの学習状況について、全国で同じタイミングで評価できるが、調査日の決定(各学校への周知)は、学校行事等の日程と調整する必要があるため、基本的に約1年半前に行っている。また、災害や感染症等の突発的な事情で調査日に全国一斉実施ができなくなった場合には、改めて数ヶ月以内に調査日を調整することが難しく、実施を見送らざるを得ない¹⁹。
- また、CBTでも現行の方式と同様に全国同日一斉で実施する場合、参加児童生徒数が約200万人と大規模であるため、サーバ及びネットワークの負荷はかなり大きい。例えばセンター集約型のネットワークの場合、一度にアクセスが集中するため、効果的な負荷分散の仕組みが必要である。さらに、校内外に通信速度が遅くなる原因(ボトルネック)があった場合、調査問題の配信時や解答の送信時等に不要な待機時間が生じるなどして、児童生徒の解答への動機づけが低下したり、高いストレスを感じさせたりする可能性があり、これらは、学力を的確に把握することを妨げる要因となってしまう。加えて、ネットワーク機器の故障や大規模な通信障害等が発生した場合には、全国一律に設定した調査日に調査自体が実施できない可能性も十分想定される。
- 今般、専門家会議において提言された悉皆調査の目的を踏まえれば、当該調査は入試等の選抜試験と基本的な性格を異にするものであり、その設計や実施にあたり、入試等と同様の厳格性を適用することは必ずしも要しない。また、学力の測定技術の発展状況も踏まえると、現行の全国同日一斉という実施方式については、CBT化を契機に、より柔軟な仕組みとすることが適当と考えられる。
- こうした技術・運営上の理由や、学校における実施日程の柔軟性の確保といったことなどを踏まえ、悉皆調査については、全国同日一斉に実施するという調査設計を見直し、一定期間内(複数日に分散)で実施することが適当である。その際は、教育委員会等の判断により、各学校において都合のよい調査日を選択できるようにすることが望ましい。
この点、具体化に当たっては、後述する分冊方式を念頭に置きつつ、調査実施にかかるオペレーションや、調査問題の数や取扱い(漏洩の抑止を含む)、各教育委員会等への結果提供の在り方等とあわせて、試行・検証も活用しながら検討を行う必要がある。

¹⁹ 全国学力・学習状況調査は、平成19年度から開始以降、平成23年度と令和2年度の2回、実施を見送っている(平成23年度は東日本大震災、令和2年度は新型コロナウイルス感染症の影響による)。

- 経年調査については、全国的な学力の状況について経年の変化を把握・分析することを目的として3年に1度程度、抽出で実施されている。また、現在、調査期間は約1か月間²⁰で、学校の都合のよい日にちを設定して調査を実施している。さらに、調査問題は、重複テスト分冊方式²¹を採用しており、その調査目的から、問題は非公開となっている。

現行の経年調査の設計は、調査実施人数や調査実施時期、調査問題の取扱いなどの観点から、CBT化を進めるための要件をより多く備えているものと考えられ、まずは、現行の調査設計を基本として、具体的な作業を進めることとしてはどうか。

②調査問題

(問題セット)

- 一定期間内（複数日に分散）で実施する場合、調査問題は複数セット用意するとともに、調査に参加する個々の児童生徒の学力を的確に評価するため、比較可能な問題セットを複数用意する必要があり、基本的にIRTの採用が必要となる。また、現行の悉皆調査では、問題は調査日当日に公開されているが、例えば、調査期間中は結果データの信頼性を維持するため非公開とし、調査期間終了後に一部公開するなどの方策が必要となる。さらに、調査期間中の問題漏洩について、必要な方策を講じることが必要となる。
- 現行の悉皆調査は、その時間的制約や、問題が1セットに限られていることから、出題される問題数及び領域が非常に限られている。複数の問題セットを用意することにより、幅広い領域からの出題が可能となる。問題公開の在り方を工夫すれば、教育指導や教育施策の改善・充実という観点から大きなメリットを発揮できるのではないかと。

(問題の在り方)

- 悉皆調査の問題構成については、今後、教育指導への活用という調査の目的を達成する観点から、教科の専門家の意見を十分に踏まえ、検討される必要がある。CBT化に伴い、一部の問題は非公開としつつ、特に学習指導要領の理念・目標・内容等に基づき、教育指導上特に重視される点や身に付けるべき力を具体的に示すメッセージとなる問題についてはできるだけ公開するといった、問題構成の工夫が考えられる。
- また、現行の悉皆調査は、大問の中にいくつかの問を設ける形で出題している。問題をプールしていく際、大問形式のままプールするのか、新たに小問独立形式としてプールするのか、問題をプールする際には等化を前提とするのか否かなど、今後具体的な出題方式について、あわせて検討する必要

²⁰ 令和3年度は、6月1日（火）～6月30日（水）の期間中、対象学校が実施可能な日時に実施している。

²¹ 重複テスト分冊方式：出題するテスト項目を複数の問題ブロックに分割し、いくつかの問題ブロックをまとめて複数の分冊を構成、さらに分冊同士の内容が一部重なるように問題ブロックを配置するテストデザイン。共通の問題項目ブロックを利用してIRTを用いた等化を行うことにより、全ての分冊を共通の尺度上で比較することが可能。

がある。

(出題方法)

- TIMSS2019 において試行的に CBT で調査を実施した際、生徒からは、イラストや動画・アニメーションなどがあると紙形式の調査よりも楽しみながら問題に取り組めるとの声があった。CBT の特性を活かして、児童生徒の意欲をより引き出せるような出題や、映像や音声を活用してより実際の学習場面に即した出題も可能となる。国立教育政策研究所において、測りたい学力をよりの確に問うことができるよう、CBT ならではの出題方法として、どのような工夫ができるかさらに調査研究を進めていくことが必要である。
- 算数・数学における作図や記号を用いて解答する調査問題について、数式のエディタ(補助ツール)を使って解答することは可能だが、児童生徒が操作に不慣れな場合、解答に時間がかかり、戸惑うことが想定され、出題方法の工夫が必要となる。
- 国語の調査問題は、現在縦書きで作成されているが、例えばインターネットからの情報等は、教科書でも横書きで記載されており、問題の種類によっては、横書きで示すことも考えられる。また、複数の文章から情報を得たり考えを形成したりする能力をみるなど、現状の課題に対応した問題が、CBT でどの程度作成可能かといった検討も、試行を通じて検証することが必要である。
- CBT 化によって、TIMSS(国際数学・理科教育動向調査)における PSI(問題解決と探究)²²のように、CBT の特性を活かした新しい問題が導入できることとなり、PBT と比べて、例えば「思考力」や「問題発見・解決能力」といった能力を測定しやすくなるのではないかとの期待もある。一方で、こうした点についての実証的な検討は十分になされていない。国立教育政策研究所での調査研究と連携を図りながら、今後の試行・検証や予備調査等において、定性的・定量的な観点から検証していく必要がある。その際、動画や問題の工夫をすることによって、出題時間が長くなることによる影響にも留意する必要がある。

(記述式)

- 専門家会議において、悉皆調査は教育指導への活用を目的としていることから、学習指導要領の理念・目標・内容等に基づき、教育指導上特に重視される点や身に付けるべき力を具体的に示すメッセージとなる問題を出題することが求められていること、また、この点は地方自治体からも強く期待されていることが指摘されており、CBT 化にあたっては、記述式問題を引き続き出題することが求められる。

²² PSI (Problem Solving and Inquiry (問題解決と探究)): TIMSS2019 から加わった新類型の問題で、研究室での実験や日常生活の文脈を模した問題を出題。児童生徒の解答に応じて、次の問が変化するという従来の筆記型調査ではできない出題方式となっている。また、問題はカラフルで動画を含んだものとなっている。

- また、記述式問題の出題にあたっては、キーボード等の操作と手書き入力の違いによって、問題の正誤状況に影響を及ぼさないか、解答時間に差が生じないかなど、試行・検証を通じて、実証的に確認していく必要がある。

③項目反応理論 (IRT)

(IRT の活用)

- 悉皆調査と経年調査を国が実施すべき調査の「二本柱」とする整理の下、悉皆調査については、①で述べたとおり、CBT 化にあたって現行の全国同日一斉に実施するという調査設計を見直し、一定期間内（複数日に分散）で実施することが適当であり、基本的に IRT の採用が必要となる。また、経年調査については、異なる時点、異なる児童生徒間の比較を行うという目的上、IRT を採用する必要性は従来と変わらない。
- IRT の具体的な活用にあたっては、調査等の目的によってその使い方が異なる点に留意する必要がある。学力調査を、①個人の学力を学習の最終段階（学年）で総括的に評価すること、②集団の学力の経年変化を測ること、③個人の学力を日々の学習の中で個別最適化した形で測ること、④個人を追跡してその学力の伸びを継続して測ること、と分類した場合、悉皆調査については、主に①の観点から、経年調査については②の観点から、どのように IRT 技術を利用するか検討する必要がある。
- IRT を採用した CBT の方式には、分冊方式、LOFT 方式、適応型、マルチステージ方式があり²³、それぞれの特徴は以下のとおりである。悉皆調査については、その目的を踏まえると、調査問題が持つメッセージ性（学習指導要領に基づき、教育指導上特に重視される点や身に付けるべき力を具体的に示す媒体としての機能）と、データを活用した教育指導の充実を図ることの両面が重視される。加えて、IRT を現段階で導入していないことも踏まえると、まずは分冊方式から採用することが適切である。
 - ・分冊方式：多数の小問群（ブロック）を複数の問題セットにあらかじめ割り付ける方式。
 - ・LOFT 方式：出題領域や問題の難易度、解答時間等が同じぐらいになるように問題セットを自動編集し、出題の直前にサーバから配布する方式。児童生徒によって出題される問題が異なる。
 - ・適応型方式：小問の正誤に応じて、次に出題する問題を変える方式。
 - ・マルチステージ方式：大問内の得点（推定値を含む。）に応じて次に出題する大問を変える方式。
- なお、IRT モデルは、多肢選択式のみならず、記述式問題にも適用可能である。記述式問題の採点結果が数値化されていれば、段階反応モデルと呼ばれる IRT モデルの適用が可能となる。

²³ 令和2年6月8日開催 第2回本 WG 寺尾委員ご発表資料

(悉皆調査における IRT)

- 現行の悉皆調査は、全国の児童生徒が同じ問題を解答し、その結果から課題を把握・分析することによって、教育現場の指導改善等に活用している。また、配点や部分点がなく、どの問題も、正答・誤答のいずれかとなり、調査結果は、正答数や正答率で表している。さらに、調査問題は、毎年度、限られた問題数の中で、全ての児童生徒に身に付けておいてほしい事項を問うており、年度間で難易度を調整しているわけではないため、基本的に、結果の年度間比較はできない仕組みとなっている。
- 悉皆調査に IRT を採用する場合、調査前に各問題の難易度や識別力を把握することが可能であるため、一斉実施ではなく、相互に比較可能な複数の問題セットが準備でき、一定期間内（複数日に分散）での実施や、方式によっては年度間の結果の比較が一定程度可能となるなどの利点が得られる。一方で、問題を継続的に使用する必要があるため、非公開とする問題が必須であることや、原則として調査前に各問題の特性（難易度や識別力等）を確認する必要があること、問題バンク²⁴を用いる方法ほどではないものの、現行よりも多くの問題を作成・準備する必要があること、問題を長期間にわたって管理することなどが大きな課題となる。この点については、調査の全体設計との整合性をとりながら、試行・検証も活用して計画的に取り組んでいく必要がある。
- 特に、悉皆調査に求められている教育指導への活用という観点からは、問題公開と結果のフィードバックは重要であり、IRT を採用する場合でも、調査問題は一律非公開とするのではなく、とりわけ教育指導上特に重視される点や身に付けるべき力を具体的に示すメッセージとなる問題については、一定数公表し、学校現場で教育指導の改善等に活用できるようにする必要がある。
- 児童生徒の解答データを統計的なモデルの下で取り扱う IRT の運用に際しては、一般に項目母数（調査問題の難易度や識別力等）の推定誤差や、等化をする段階での推定誤差の大きさに留意が必要である。悉皆調査で IRT を採用する場合、過去に実施された調査の解答データを用いて等化する方法も考えられる。その運用が長期になると、運用開始時にはわずかだった誤差が累積して大きなものになり、その結果として学力についての議論の的確性を阻害する可能性がある。限られたデータと項目数で運用を行う以上やむを得ないこととも考えられるが、こうした点に留意しつつ、長期的な運用を見越した等化の方法やアンカー項目数の設定等について、技術的観点から検討する必要がある。

(経年調査における IRT)

- 経年調査は、全国的な学力の状況について経年の変化を把握・分析し、教育施策の検証・改善に役立てるという目的から、IRT 技術として、重複テスト分冊方式を採用している。まずはその方式で CBT 化することとし、将来的には、さらに効率的な測定方法について、PISA 等の国際学力調査にお

²⁴ 問題バンク：調査問題そのものと、各問題の難易度や識別力等を蓄積したデータベースのこと。調査問題の属性に関する情報（メタ情報）として、出題回数や出題内容・領域、解答時間等を追加する場合もある。

ける取組も参考にしつつ、専門的・技術的検討を進めることを期待したい。

(作問体制)

- CBT ではPBT ではできなかった様々な出題が可能になるが、作問にあたっては、教科教育の専門家に加え、CBT での出題方法や解答形式等に詳しい専門人材が必要になるため、作問体制の在り方を見直し、相当強化する必要がある。
- IRT を採用した調査問題を作成する場合は、大まかに以下の役割が必要となる²⁵。こうした役割を踏まえた上で、作問を担当する国立教育政策研究所において、CBT 化に向けた体制を構築していく必要がある。
 - ①管理者：IRT や教育全般に関する一定の知識を有し、開発する調査とそれに必要な調査問題セットまたは問題バンクの構築・維持管理を行う者
 - ②編集者：IRT と教科・教育に関して一定の知識を有し、かつ、項目やテスト冊子の編集に必要な技術と感性を持つ者
 - ③作成者：教科内容に精通している者
 - ④分析者：テスト理論に関して専門的な知識と技能を持つ者

(問題漏洩への対応)

- IRT を採用する場合、一定数の問題を継続的に使用する必要があるため、問題漏洩への対応が必要となる。漏洩の発生原因や調査設計によっては、漏洩が起こったこと自体の把握が困難であるという可能性を踏まえ、予防の観点からも対策を検討することが重要である。例えば、医療系大学間共用試験では、受験者を含む試験関係者から守秘に関する同意を得たり、システム上もスクリーンショットや印刷ができないようにしたりするなど様々な手段を講じている。また、正答率等の指標値を確認し、仮に大幅に数値が変化した場合は、その後の出題候補から削除するなどの対応をとっており、全国学力・学習状況調査についても、こうした取組を参考とすることが必要である。
- また、特に悉皆調査においては、一定期間内で複数日に分散して実施することとなると、問題漏洩は起こり得ることとして対応を準備しておく必要がある。仮に問題セットごと漏洩した場合、調査の実施に多大な影響が生じることから、人為的なものであれば、漏洩させた者に事後的にその責任を問う仕組みも検討しておく必要があるのではないか。
- 個別具体的な問題漏洩対策に止まらず、不適切な行為の背景・誘因となる関係者の意識の在り方に目を向けることも重要である。悉皆調査は、児童生徒の学力の状況を把握・分析し、教育施策の成果と課題を検証し、その改善を図るための、あくまで「調査」であり、地方自治体や学校を競争させた

²⁵ 令和3年4月23日開催 第9回本WG別府正彦氏ご発表資料

り、序列化したりするためのものではないということが社会全体で広く共有されていくよう、調査の趣旨について更なる周知徹底を図っていくことが望まれる。

④結果提供・解答データの取扱い

(結果提供)

- 現行の悉皆調査は、学校における指導改善等に活用する観点から、全ての調査問題を公開している。これを全て非公開とする場合、結果が提供された後、児童生徒が問題そのものを見ることなく、数値だけをもって自分の学習の、どこをどのように変えていく必要があるかを理解することは困難である。また、問題そのものを通じて教師に対する指導改善のメッセージを発するという役割・機能も失われることになる。大学生を対象とした資格試験である医療系大学間共用試験と異なり、悉皆調査については、児童生徒の学習改善に資する調査目的を達成するため、一定数の問題公表を行うことを前提に検討を進める必要がある。
- 公開問題の解答に対するフィードバックを行う方法の在り方に加え、非公開問題についても、その問題が測ろうとする領域や資質・能力が明確になっていれば、個々の問題内容に具体的に触れなくても、児童生徒の得手不得手についてフィードバックすることは可能ではないか。
- 現行の悉皆調査では、各教育委員会、学校及び児童生徒に結果を返却しているが、特に児童生徒へ返却する内容について、学習改善に活用することを目的とする調査であることを踏まえ、日々の学習で活用できるものであることが必要である。IRTを採用した場合、問題の難易度に関する情報もあわせて返却することにより、児童生徒の解答状況と照らして、より児童生徒の理解度に応じた効果的な復習内容や宿題を提示することなどが可能となるのではないか。
- これらの点を含め、各教育委員会や学校、児童生徒への結果提供や、結果公表の在り方については、海外の学力調査における方法を参考にしつつ、専門家会議等において教育委員会等の意見も聴取するなどして、さらに詳細に検討を行う必要がある。

(解答データの取扱い)

- CBTでは、解答に加えてログ²⁶も把握することによって、児童生徒の解答中のつまずきをはじめとした多面的な分析が可能となり、指導改善にもメリットがある。また、ログは、児童生徒の解答状況や学力のありようを細かに把握し、指導改善に活用するための基礎的な研究資料にもなり得ることから、その重要性は高い。なお、それらの分析にあたっては、データの前処理や分析等を行うことができる専門的な技術を持った人材・組織が不可欠となる。

²⁶ ログ：児童生徒が解答までに要した時間やキーストローク、クリックなどの動きなどの操作に関する詳細データ。

- 解答データについては様々な利用・分析が可能になると考えられるが、解答データ等の取扱いやセキュリティについても十分に検討しておく必要がある。
- 悉皆調査は、全国的な児童生徒の学力や学習状況を把握・分析することと同時に、学校における個々の児童生徒への教育指導や学習状況の改善・充実等に役立てることも重要であり、その点を踏まえた分析の在り方について検討する必要がある。

⑤特別な配慮が必要な児童生徒への対応

(CBT システムにおける合理的配慮)

- 障害のある児童生徒数は近年増加傾向にあり、特別な配慮が必要な児童生徒は、特別支援学校や特別支援学級、通級指導教室のみならず、通常の学級にも在籍している。CBT を設計する際は、特別な配慮を必要とする児童生徒が全ての学校・学級に在籍することを前提として考えることが必要不可欠である。また、療養中の児童生徒及び障害のため通学して教育を受けることが困難な児童生徒が、自宅や病院等で調査を受ける可能性があることを最初から想定しておくことが重要である。
- デジタル教科書や教材に搭載されている代表的なアクセシビリティ機能は、調査の実施に関して以下のように整理できる。この他、例えば、聴覚障害のある児童生徒については、画面上で様々な指示を表示できるようにするなど、CBT システム上、可能な限り合理的な配慮を行い、本人の意向や普段の学習環境にあわせて対応できるようにする必要がある。
 - ①問題表示：文字や図形をハイライトしながらの音声読み上げ、ピンチ操作による文字や図形等の拡大・縮小、配色の変更、ページジャンプ、文字サイズを変更しても上下のスクロールだけで文章にアクセスできるリフロー表示、ルビのオン・オフ、フォント・行間隔・文字間隔・縦書きや横書きなど、文書スタイルの変更
 - ②問題の解答：キーボードや外部スイッチなどから操作できる代替入力
 - ③補助機能：ブックマーク（付箋）添付機能、メモやラインマークなどの書き込み機能
- タッチディスプレイや音声入力等については、端末ごとに機能が異なるため、当該児童生徒が普段使用しているデバイスでの調査を実施できるようにしてはどうか。また CBT システム上、どの程度の機能を実装するのか検討する必要がある。

(作問・出題における合理的配慮)

- 多様な特性を有する児童生徒に対応できるよう、出題・解答のインターフェイス²⁷でどのような工夫が可能になるのか、試行・検証等を通じて解答時間や解答事例等も集積しつつ、検討していく必要

²⁷ インターフェイス (interface) : ここではユーザインターフェイスのことを指し、CBT で出題された問題に対して児童生徒が解答するときの操作画面や表示内容等の総体。

がある。

- 作問・出題にあたっては、色に関する配慮が必要となる。紙での出題の場合は、色の使用が制限されるのでそれほど差異は生じないが、CBTになるとカラフル素材等多くの色が使用可能となるため、どのような配慮や工夫をする必要があるか、検討する必要がある。例えば、図や写真等の視覚的な表現や発音等の聴覚的な表現だけに依存しない問題を作成することが重要である。
- 適応型方式の場合、例えば視覚障害者や聴覚障害者への出題が困難な問題について、児童生徒の障害の特性に応じて、問題バンクから難易度や問う内容が一致する他の問題に自動的に代替することも可能となるのではないかと。

(3) CBT 特有の課題・論点

①問題作成の体制や工程

(問題作成の体制)

- 現行の悉皆調査の問題は、国立教育政策研究所に作問委員会を設置し、約1年かけて、学習指導要領の理念・目標・内容等に基づき、教育指導上特に重視される点や身に付けるべき力を具体的に示すメッセージとなるよう専門家による検討を重ね、入念に作成されている。CBT化に伴う技術的対応に加え、IRTの採用にあたっては、冊子の数や公開・非公開等の問題の種類・構成をはじめ、従来の作問の在り方は大きく変わることになるため、現在の作問体制を見直し、強化する必要がある。
- 悉皆調査にIRTを採用する場合、調査に含める問題項目や具体的に出題する問題項目の選択をどうするかも新たな論点となる。問題項目の選択については、教科の専門性の観点と技術的観点の双方が必要である。国立教育政策研究所においては、令和3年10月に「教育データサイエンスセンター」が新設される予定であるが、これを機として、テスト理論やデータ分析に精通している人材を一定数配置し、専門の組織を設けるなど、検討・実施体制を強化することが不可欠である。また、CBT化にあたっては、問題開発と開発した問題を整理・分類するためのコードの付与やその編集・管理を行う体制、すなわち、問題の質の保証と技術的な支援を組織的に行うための体制も不可欠である。

(問題作成の工程)

- 例えば、医療系大学間共用試験では、各大学から広く様々な問題を集め、精査、吟味していくというプロセスで作問を行い、問題バンクを構築している。一方、一部の地方自治体で行っている学力調査の問題傾向を見ると、悉皆調査の問題を作問の基準として、様々な工夫をして出題しているものが多い。今後、こうした実例や経緯を踏まえつつ、悉皆調査のCBT化に伴って想定される作問数の増大に対応するため、大学や教育委員会等との連携を視野に入れながら合理的な工程の在り方を検討していく必要がある。

- 経年調査については、その調査目的から、基本的に毎回同じ問題を用いているが、現在の PBT による調査問題が、CBT 化した場合でも引き続き使用可能な問題であるか、これまで PBT で測定してきた過年度の調査における学力との同等性や過年度の結果との比較可能性の観点から精査を行いつつ、CBT 化に向けた取組を進めることが必要である。

② 調査資材の印刷、配送・回収、採点、集計、分析等、現在の一連の工程の効率化

(問題冊子の印刷、配送・回収)

- CBT 化により、調査資材の印刷作成部数の大幅な削減が期待され、印刷コストの低減とともに、配送・回収に必要な箱数の減少や小型化等によるコスト低減が期待される。また、これら大量の調査資材を扱うための準備期間が3か月程度短縮できる可能性がある。なお、現在は、調査日の前日までに配送を行い、調査日の翌日に回収を行っているが、完全な CBT 化が実現できれば、その配送・回収日の制約も解消されることになる。
- 一方で、コストに関しては、新たにネットワークインフラやシステム関連の ICT 環境の整備、中間サーバを活用する場合にはその保守・点検等にかかる経費等が、相当程度必要となることに留意する必要がある。

(採点)

- 現在、各教科及び質問紙調査の解答（回答）用紙は OMR²⁸でデータ化しており、その読み取り枚数は、小中学校それぞれ約 100 万枚必要となる。CBT 化により、その読み取り枚数や読み取りにかかる期間及びコストの大幅な削減が期待される。
- 選択式問題については、CBT 化により、自動採点が可能となる。また、短答式問題の採点については、現在、解答用紙にある解答欄を OMR で読み取った上で、システムによりテキスト化し、文字列ごとに人が目視確認しているが、その業務にかかるコスト削減が期待される。加えて、システムが判読できなかった解答の目視採点にかかるコスト削減も期待される。
- 記述式問題の採点については、現在、解答用紙にある解答欄を OMR で読み取った上で、人が目視採点を行っている。解答自体がデータ化されることによって、OMR で解答用紙を読み取るコストや、判読困難な解答の採点にかかるコストの削減が期待される。
- また、現在、理化学研究所による AI を活用した記述式問題の自動採点技術の研究も進められている²⁹。記述式問題にかかる結果のフィードバックの迅速化や、人による採点にかかるコストや時間の

²⁸ OMR (Optical Mark Reader) : マークシート用紙を読み取る光学式マーク読み取り装置のこと。選択式問題の試験やアンケート調査等大量の情報を処理する際に使用される。

²⁹ 令和3年1月26日開催 第7回本 WG 乾健太郎氏 (理化学研究所 A I P センター自然言語理解チームリーダー) ご発表資料

圧縮を図るという観点から、将来的な実装を見据えて、継続的に調査研究を行っていくこととしてはどうか。まずは、人と機械のよりよい役割分担や自動採点技術の可能性について、MEXCBTにおいて試行・検証を行い、そのフィージビリティを高めていけるよう取り組んでいくこととしてはどうか。

(結果提供)

- 現在、集計終了後、専用の Web システムを通じて、教育委員会や学校に納品帳票、ローデータ、集計支援ツールなどを提供しているが、自動採点で即時採点が可能な問題構成であれば、調査終了後、速やかに調査結果や集計結果を提供することが可能となる。
- CBT 化した場合、学校現場では、紙の調査資材の受け取りや回収の代わりに、端末やネットワーク状況の確認作業等が生じることが考えられ、CBT による実施でも、一定の負担はありとされる。一方で、教育指導に必要な資料が速やかにフィードバックされることとなれば、教師の指導改善や児童生徒の学習改善に早期に繋げる観点からは有益である。

③学校現場における円滑な実施

(学校現場への支援)

- 平成 31 年度に CBT により実施された中学校英語「話すこと」調査の「検証報告書」³⁰において、教職員の調査の準備・実施にかかる作業への負担感について指摘されている。例えば、前年度に事前検証を行い、新年度当初に調査プログラムの展開等の準備を行うスケジュールについて、教職員の人事異動や新年度の校務分掌決定時期を挟み、時間的・心理的余裕がない中での準備であったこと、学校によっては、調査プログラムの展開や音声データの回収に数時間要し、負担が大きかったことが指摘されている。
- また、設置管理者からも、学校現場への負担にかんがみ「小規模での実証・改善を段階的に重ねることが必要」との指摘もなされている³¹。当日調査を担当する教職員の心理的負担をできるだけ軽減する必要性も十分に踏まえ、調査の準備、実施中、終了後に至るまで、学校や設置管理者、教職員が円滑に調査を実施できるよう十分な配慮が必要である。例えば、端末等のトラブルが発生した場合、教職員がスムーズに対応できるよう、教職員向けのマニュアルなどの作成や基本的な対処方法を学ぶ機会を設けるなど、実施にあたってのサポートが必要である。
- PISA 調査の日本での実施においては、学校で調査を円滑に実施するため、調査実施に向けた説明会の開催や、対象校向けにコンピュータを使用した演習を体験できる機会を設けること、動画を用いて実際の調査実施手順を紹介するなどの対応を行っている。こうした取組も参考にしつつ、令和 3 年

³⁰ 平成 31 年度（令和元年度）全国学力・学習状況調査中学校英語「話すこと」調査検証報告書（令和元年 9 月 20 日 全国的な学力調査に関する専門家会議平成 31 年度英語「話すこと」調査検証ワーキンググループ）

³¹ 令和 3 年 3 月 30 日開催 第 8 回本 WG 埼玉県教育委員会ご発表資料

度以降の試行・検証を活用しながら、学校への具体的な支援の在り方について検討を行っていく必要がある。

(4) 実施体制等

① CBT システムの開発（業務管理、採点、集計・分析、問題バンクなど）

(システム設計)

- システム設計は、どの程度の技術・サービス水準を要求するかを明確にする必要がある。例えば、実施方式に関するシステム、問題作成や管理に関するシステム、解答データの回収、採点、集計・分析等に関するシステムなど、様々なシステムの技術開発を同時並行で行おうとすると、必要な予算の確保が困難な状況も予想されるため、優先順位を定めて設計・開発する必要がある。また、地方自治体や学校によって、整備されている端末や導入されているアプリケーションなどが異なるので、様々な OS やブラウザに対応できるように設計する必要がある。
- また、CBT システムには大規模調査特有の機能を搭載する必要がある。主には、調査が適切に実施されているかを確認する機能（いわゆる「監督機能」）を持たせることが必要となる。監督機能には様々なモードがあるが³²、全国学力・学習状況調査の場合、解答（回答）する者が児童生徒本人であることの確認や、児童生徒が調査を適切に実施し終えたかどうかを教職員が確認するモードを搭載する必要がある。また、調査問題が正しく配信されたことを確認する機能や、解答（回答）が確実に送信されたことを確認する機能も搭載する必要がある。
- 現行では、毎年度入札で委託事業者を決定しているが、CBT システムに関しては継続的なシステム運用が不可欠であり、複数年度にかかる契約の仕方や、CBT システムの管理・運用に関する国と事業者の役割分担の在り方、運用にかかるノウハウを国においてどう蓄積していくかといった点について、MEXCBT のシステム運用とあわせて検討していく必要があるのではないかと。

(実施方式)

- CBT の実施方式としては、主に①USB 方式³³、②中間サーバ+LAN 接続方式³⁴、③WAN 接続方式³⁵がある。全国学力・学習状況調査を CBT 化する場合、①USB 方式は、大量の USB が必要となるため経費が必要となることや、学校現場での取扱いにも手間がかかり、調査実施の効率化には直結しない。②中間サーバ+LAN 接続方式は、経費（初期費用・維持費）と実施にかかる負担のバランスなどの観点から現実的と考えられるが、中間サーバの数によっては経費が膨大となる。③WAN

³² 令和2年6月8日開催 第2回本 WG 寺尾委員ご発表資料

³³ USB に調査アプリを入れておき、ネットワークを利用せずに実施する方式。

³⁴ サーバを教育委員会や学校に設置し実施する方式。児童生徒の受験端末は、調査実施中は域内ネットワーク、校内 LAN にのみ接続し、外部ネットワークには接続しない。解答は一時的に中間サーバに送信しておき、解答終了後に時間差でメインサーバに送信する方式。

³⁵ 直接メインサーバに接続する方式。

接続方式は、ネットワークが不安定な状況ではリスクが高いことなどが課題となるが、今後、GIGA スクール構想の推進によりネットワーク環境の改善・整備が進められることが見込まれる³⁶。

こうしたことを踏まえ、学力調査が的確かつ着実に実施できる基盤を維持できるよう念入りに設計を行うことを前提として、③を基本としつつ、地域のネットワーク環境等を考慮し、②も組み合わせて、試行・検証することとしてはどうか。

- 通信環境等のインフラを誰が維持・管理するのかという点について、校内通信ネットワークの整備は文部科学省の GIGA スクール構想で進めているところだが、実際の個々の学校における通信環境の維持・管理は、個々の地方自治体等が主体となって行うことを基本として検討すべきである。

(問題バンクの作成)

- 問題バンクについては、事前に数万題もの問題を用意する事例もあるが、どの程度の規模の問題バンクを用意するかはケースバイケースであり、調査の実施目的に応じて計画を策定し、最適規模の問題バンクを構築し、過剰な経費をかけないことが必要である。なお、経年調査では、重複テスト分冊方式を採用しており、いわゆる問題バンクは作成していない。

(地方自治体が独自に実施する学力調査との連携)

- 地方自治体自らが教育施策の検証等に活用することを目的として独自に実施している調査について、地方自治体から国との連携について提案がなされている³⁷。また、例えば京都府や埼玉県においては CBT による学力調査の実施に向けて取り組んでおり、全国学力・学習状況調査の CBT 化に伴い、どのような連携を図ることができるのか、専門家会議においてさらに検討が行われる必要がある。また、ハード面において、CBT システムについては、MEXCBT を活用できるようにするなど、一定の連携を図ることについて検討してはどうか。

②試行・検証、予備調査、試行等を含めたスケジュールや具体的な進め方

(試行・検証の必要性)

- 全国学力・学習状況調査のような大規模調査を CBT 化する場合、当初から設計しておくべき点と試行・検証を通して改善できる点を区別して進めることが重要である。後者に関しては、調査研究や小規模からの試行を通じて、その過程で得られた課題から大規模調査で実現可能な方法を検討していく必要がある。あらかじめ CBT システムの完全な要件定義を決めることは難しく、技術の進展状況も踏まえつつ、試行・検証を積み重ねる中で、新たなニーズの把握や軌道修正等も行いながら構築し

³⁶ GIGA スクール構想により整備される 1 人 1 台端末については、この端末からネットワークを通じてクラウドにアクセスし、クラウド上のデータ、各種サービスを活用することを前提としている。このため、学校内のみならず学校外と繋ぐネットワークが高速大容量であること、自治体等の学校の設置者が整備する教育情報セキュリティポリシーなどにおいて、クラウドの活用を禁止せず、必要なセキュリティ対策を講じた上でその活用を進めることが必要である。

³⁷ 令和 3 年 3 月 22 日開催「第 8 回全国的な学力調査に関する専門家会議」京都府教育委員会ご発表資料

ていくことが重要である。

- 文部科学省が令和3年度から実施する CBT 化に向けた試行・検証では、MEXCBT を CBT のプロトタイプシステムとして活用することとしている。同システムを活用した試行・検証を通じて、全国学力・学習状況調査のような大規模調査を実施するために必要な機能について精査し、実装に向けた準備を行っていく必要がある。
- なお、試行・検証にあたっては、実施体制や調査理念の周知を行った上で、各ステークホルダー（教育委員会、学校、教職員等）が参加しながら開発を進めていくことが望ましい。

（準備期間と進め方）

- 全面実施までの準備期間について、例えば国際学力調査の場合、PISA は、2003 年に「CBT 化の計画」を決定し、2006 年の予備調査、2009 年に本調査における一部の分野から CBT を導入し、2015 年に全面実施という経過をたどっている。また、TIMSS については、2014 年に CBT の開発計画が公表され、2019 年は PBT と CBT の選択制で実施、2023 年に全面実施というスケジュールで進めている。いずれも、PBT との併用期間を含めて全面実施に至るまでには相応の準備期間を要し、また、一部の分野や対象から段階的に CBT 化を進め、全面実施につなげている。また、医療系大学間共用試験については、2002 年から「トライアル 0」としてプロトタイプシステムの開発を開始し、「トライアル 4」まで段階的に試行を実施し、2005 年に CBT 化の実現につなげている。
- 全国学力・学習状況調査については、GIGA スクール構想によって各学校に配備された 1 人 1 台端末を活用してオンラインで実施することや、地方自治体や学校によってネットワーク環境や整備されている端末、導入されているアプリケーションなどが異なること、また、特に悉皆調査はその対象が小学校と中学校あわせて約 3 万校、児童生徒数は約 200 万人という大規模であり、それに応じたシステムなど必要な実施体制を準備する必要がある。また現行よりも多くの問題が必要となることも踏まえると、CBT への移行にあたっては、相応の準備期間と費用を要する。今後、令和3年度の試行・検証事業を皮切りとして、課題の抽出とその解決を繰り返しつつ、段階的にその規模・内容を拡充させながら、全国レベルでの CBT 化の実現に着実に繋げていくことが適当である。

（CBT 化に向けた「工程に関する考え方」）

- これまでの整理を踏まえ、全国学力・学習状況調査の CBT 化については、段階的にその規模・内容を拡充させながら進めることとし、令和6年度に予定されている次回の経年調査から順次導入することとしてはどうか。

経年調査は抽出であり、悉皆調査より対象学校数及び児童生徒数が少ないこと³⁸や調査期間は約 1

³⁸ 経年変化分析調査の対象学校数は約 1,350 校、対象児童生徒数は約 11 万人（令和3年度）。

か月間であること、調査問題は国全体の学力状況の経年の変化を把握することを目的としており、基本的に非公開で継続的に使用しているものであることなどを踏まえると、CBT化に向けた要件をより多く備えているものと考えられる。また、全国学力・学習状況調査全体のCBT化を進めるにあたって、技術開発の中間目標としても位置づけられるのではないかと。

- なお、これまでPBTで測定してきた過年度の調査における学力との同等性や過年度の結果との比較可能性を担保するため、CBTで実施する初期段階、少なくとも次回の調査においては、CBTとPBTを併用した形で実施することが必要である。また、これを悉皆調査における教科調査のCBT化に向けた大規模実証の一つとしても位置づけ、システムやネットワークなどの実証や調査実施の業務フロー、人員配置・役割分担の検証、課題の洗い出しを行うとともに、調査方式による解答への影響（モードエフェクト）等を検証する機会として活用することとしてはどうか。
- 悉皆調査については、まず、児童生徒質問紙調査のオンラインによる回答方式について、次回の経年調査と同時期（令和6年度予定）を目途に、原則として全児童生徒を対象に導入することを目指してはどうか。

すでに令和3年度の悉皆調査において、一部の国立大学附属学校で端末を利用したオンラインによる回答方式を試行的に実施していること³⁹、また、児童生徒質問紙調査は選択肢式のみアンケート形式であり、システム・ネットワーク面及び児童生徒の端末の操作等にかかる課題は教科調査よりもクリアしやすいと考えられるため、教科調査に先駆けて全面導入することが可能と考えられる。なお、児童生徒質問紙調査とともに毎回実施されている学校質問紙調査については、すでに平成28年度から学校の端末を用いたオンラインによる回答方式で実施されている。
- 悉皆調査における教科調査については、経年調査と比べて実施規模が格段に大きいため、上述のように、ネットワーク・システム面の検証が必要である。また、主に指導改善に活用する目的を踏まえ、問題作成や公開の在り方、結果分析・提供等について、十分な準備や検討が必要となる。さらに、文部科学省及び国立教育政策研究所においては、毎年度の悉皆調査と並行してCBT化に向けた諸準備を進めることも必要となる。これらのことと令和3年度から実施する試行・検証や、上記次回の経年調査のCBTによる実施結果等も踏まえ、次回の経年調査の翌年度以降できるだけ速やかに、まず中学校から導入（PBTと経過的に併用）することを目指してはどうか。
- 中学校を先行させる理由としては、①情報活用能力調査の結果を踏まえると、基本的な端末の操作が小学校段階よりも熟達しており、短答式・記述式問題にも比較的対応しやすいと考えられること、②国及び教育委員会等の実施体制の面からも、まずは学校数が少ない中学校において円滑に実施できるよう準備を進めていくことが適当であることなどが挙げられる。

³⁹ 国立大学附属学校の一部（約100校、児童生徒数は約1万人）において試行的に実施。

○ なお、中学校英語「話すこと」調査は、テキストデータよりもデータ量が格段に大きい解答音声データを取扱うことや、学校における録音環境の事前確認をどのように行うかなど、CBT方式の中でも技術的に特に難易度が高い。平成31（令和元）年度にCBT（USB方式）で実施された際、欠損データが発生したことや、音声データの回収等に手間がかかり学校現場への負担が大きかったとの指摘も十分に踏まえ、慎重に検討・準備を進めていくことが必要である。

これらを踏まえ、次回の中学校英語「話すこと」調査については、全国学力・学習状況調査全体のCBT化に向けた工程の一部として位置づけた上で、教育委員会等の意見も踏まえつつ、技術的に可能な範囲で実施することが適当である。

○ 小学校については、児童の基本的な端末の操作にかかる状況を試行・検証や情報活用能力調査等によって継続的に把握しつつ、中学校におけるCBT導入の翌年度以降、できるだけ速やかに導入（PBTと経過的に併用）することを目指してはどうか。

なお、小学校、中学校ともに、CBTを導入する初年度の規模については、令和3年度以降、継続的に実施する試行・検証及び経年調査を活用した大規模実証の結果等を踏まえ、検討する必要がある。

○ 今後、本WGでは、この「工程に関する考え方」を踏まえつつ、文部科学省及び国立教育政策研究所等から、試行・検証や学力アセスメントにかかる調査研究の状況、詳細な調査設計に関する検討状況等について適宜聴取しつつ、CBT化に向けた工程の進捗管理を行うこととする。

なお、現時点では、技術発展の可能性や試行・検証を経て明らかになることなど、不確定な要素が様々あるため、それらを踏まえ、今後、ここに示す目標時期や手順等を変更することもあり得る。その際は、本WGにおいて状況や課題等を聴取の上、その解決に向けた方策や変更の方向性等について、専門的・技術的観点から検討を行い、CBT化の実現に向けた取組が着実に進められるよう、適宜助言を行うこととしたい。

○ 本最終まとめの内容は、現在の技術的制約・条件の下で、全国学力・学習状況調査のCBT化について専門的・技術的観点から検討を行ったものである。現在、政府においては、デジタル庁の創設等、デジタル化が推進されているところであり、個々の児童生徒への教育指導の改善等を目的として現行全国同日一斉で実施されている悉皆調査や、全国的な学力の状況について経年の変化を把握する経年調査が、将来的な技術進展の可能性を見据えつつ、今後、より良く効率的な形で実施できるよう、必要なネットワーク環境の整備・改善等の取組がより一層進められることを強く期待したい。