

全国的な学力調査に関する専門家会議
「全国的な学力調査の CBT 化検討ワーキンググループ」
主な意見（概要）

第 1 回：令和 2 年 5 月 21 日

第 2 回：令和 2 年 6 月 8 日

※議事概要をもとに事務局作成

<CBT 化のメリット>

- CBT 化した場合、紙による筆記方式に比べて、解答、採点、データ化への流れが合理化・効率化される。仮にオンライン化が実現すれば、選択式の問題は自動採点が可能となり、実施から結果提供までの期間が短くなるため、教育委員会や学校における指導改善を加速・向上することが期待される。
- 現在は、調査日に登校している児童生徒しか調査を受けることができないが、CBT 化により（家庭のパソコンも使用できる場合）、例えば、不登校の児童生徒も調査に参加できるようになるなど、こうした児童生徒の学力や学習状況を把握し、指導に活かすことが可能となる。
- 解答のログ等をとることにより、児童生徒の解答（結果）だけではなく、思考（過程）も可視化できる可能性がある。
- TIMSS（国際数学・理科教育動向調査）における PSI(問題解決と探求)のように、CBT 特有の新しい調査問題が導入でき、紙による筆記方式では取得困難なデータを得られる可能性がある。
- CBT 化は、紙による筆記方式に比べて、個別最適化学習のツールになりうると考えており、現在の文部科学省の目指す教育の方向性と合致しているのではないかと考えている。

<CBT 化にあたり留意すべき点>

【全体的なことについて】

- 大前提として、児童生徒が日常的に ICT 機器を使用した学習に慣れることが重要であり、その延長に、学力の把握・測定の CBT 化があると考えるべきである。
- 情報活用能力という新たな能力が学力の把握・測定に影響することとなるが、ICT の整備状況や普段の利用状況が児童生徒によって様々であり、ICT 環境の格差が広がっていると認識している。これらについては、全国学力・学習状況調査の本格的な CBT 化の前

に、予備調査を行い、測定したい学力と情報活用能力などの影響について分析することが望ましいのではないかと。

- 全国学力・学習状況調査を CBT 化する場合、現在のように、全国の小 6 及び中 3 の児童生徒（約 200 万人）が実施する場合には、（TIMSS のような数十校の抽出調査とは異なり）大規模なシステムを開発・整備することとなり、相応の費用や準備期間が必要となる。CBT 化した学力調査の実行可能性については、しっかりと検討する必要がある。
- 紙による筆記方式とは異なる想定外のトラブルが生じる可能性があり、その対応について検討しておく必要がある。

【調査問題や作問について】

- CBT 化する場合、紙による筆記方式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、現行のように教科に関する専門的な知識を持つ人間だけではなく、CBT の出題方法や解答形式などの技術に詳しい人材が必要になるのではないかと。
- CBT 化によって、算数・数学の作図など、一部の分野については、出題や解答方法に工夫を要することになるのではないかと。
- 国語に関しては、例えば、問題文を縦書きのスクロールとした場合、児童生徒は違和感を抱かないだろうか。漢字、古文や漢文などの表記についても、どのような取扱いとするか検討する必要がある。また、日本のテスト文化の中で根付いてきた縦線、波線、二重線といった表記なども、CBT ではシステム上難しい。今後、CBT に適した新たな出題形式等を検討することが重要であると考える。
- CBT 化した場合、記述式問題をどのように扱うか、よく検討する必要がある。キーボード入力とするのか手書き入力とするのか。キーボードによる解答は、一部の児童生徒には負担・困難だと思われる。例えば、記述式問題の解答は紙による筆記方式とするなど、対応を検討すべきである。
- 記述式問題は自動採点が技術的に難しいため、採点が早くなるとは限らないのではないかと。
- 解答のログ等を分析するにあたっては、データの前処理や解答時刻の分析を行うことができる専門的な技術を持った人材が不可欠となる。

【項目反応理論 (Item Response Theory) について】

- 全国学力・学習状況調査を CBT 化する場合、調査の規模 (約 200 万人) を考慮すると、IRT の導入を前提として検討すべきではないか。IRT を導入しない場合、実施方法や能力測定の方法などで CBT のメリットが生かしきれない可能性があるのではないか。
- 小規模の学力調査では IRT を導入せずに 1 つの問題フォームで一斉実施している事例などもあるが、大規模調査になると、サーバの負荷の観点などから同時に実施できる人数が制限されるため、分散実施することが必要になると考えられるが、その場合、複数の問題フォームの用意や問題管理の観点から IRT を導入することは有力な選択肢だと考える。
- 例えば、IRT を導入すると大量の問題を継続的に使用する必要があるが、これらをどのようなシステムで管理・運用するのかについても、よく検討すべきである。
- CBT 化した後の将来的な学習をイメージした際、IRT ではカバーしきれないような学力を測るための問題なども開発されるだろう。現状では、BKT(Bayesian Knowledge Tracing Model)や Deep-IRT などの新たな技術が開発されているが、まだ実用段階には至っていない。CBT 化は、現時点でできる部分から徐々に進めていくのがよいのではないか。
※BKT(Bayesian Knowledge Tracing Model)や Deep-IRT は、ログや要した時間などの解答にかかる時系列データから解答者の理解度を測定する新しい技術のこと。
- 本来的には、学力調査の CBT 化という実施方法と IRT を導入するか否かという点は、別々に整理すべき論点ではないかと考える。

<現在の全国学力・学習状況調査との関係>

- 全国学力・学習状況調査の目的との関係は、よく整理する必要がある。その上で、CBT 化を検討すべきだと考える。
- 調査の理念や目的と CBT の実証的な研究・試行は、並行して議論していく必要があるのではないか。様々な規模の試行を綿密に計画し、実施していく中で、CBT の実現可能性が明らかになる。運用面でどのような調査が可能なのかという点も踏まえて、調査の在り方や目的も、継続的に検討していく必要があると考える。
- 学力の測定をより正確に行うためには、例えば、IRT (項目反応理論) 等を導入して、同一問題を継続的に使用し、経年 (年度間) 比較ができるようにすることも考えられるが、その場合、調査問題は原則として非公開とになる。

- 指導改善に活かすのであれば問題公開と結果のフィードバックは重要である。IRT（項目反応理論）等を導入する場合も、調査問題は一律非公開とするのではなく、何らかの工夫になるのではないか。
- 全国学力・学習状況調査は、これまで、主として、知識・技能を問う A 問題、活用を問う B 問題があり、平成 31 年度（令和元年度）調査で、知識・技能や活用を一体的に問う形式（AB 一体化）となったが、CBT 化に伴い、例えば知識・技能を問う問題を CBT 化するという仕組みも考えられるのではないか。

<CBT システム等の開発について>

- システム設計や開発は、限られた予算の範囲内で技術的水準をどの程度を求めるのかを検討しなければならない。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、集計など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、予算を案分することになるため、何を優先すべきかを決めて設計・開発する必要があるのではないか。
- 全国学力・学習状況調査の CBT 化というプロジェクトは、かつてないクリエイティブなプロジェクトだと認識している。そのためには、システム開発には、児童生徒、教師、保護者などのユーザーがどのような期待や要望を持っているかについてマーケティングした上で要件定義することが望ましいのではないか。
- 全国学力・学習状況調査を CBT 化するにあたっては、中間サーバ方式が現実的ではないかと考える。WAN 方式の場合、国がサーバを用意するため多額の経費がかかり、USB 方式の場合、大量の USB が必要となるため経費がかかり取扱いにも手間がかかることが懸念される。
- CBT 化した場合、解答データについても様々な利用・分析が可能になると考えられるが、他方、データの読み取り機能もかなり向上しており、データを公開するだけで個人情報復元できるという可能性があるという側面もある。CBT で得られる解答データの取扱いやセキュリティ等についても、十分に検討しておく必要があるのではないか。
- 実際に CBT システムを開発した場合、事業者とどのような契約で実施・運用するのかという点は検討を要するのではないか。例えば、現行調査では、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、国と事業者の役割分担等について整理すべきだと考える。

<本格実施に至るまでのスケジュール等について>

- 全国学力・学習状況調査を CBT 化する場合、最初から大規模で実施しようと思うと、システム開発に関して技術的にも予算的にも厳しくなり、またエラーやトラブルも生じる可能性が高いと考えられる。そのため、試行や予備調査については、抽出などで小規模で行っていくことが望ましいのではないか。これらを経て、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と考える。