

令和4年度 教育研究公開シンポジウム報告書

初等中等教育—052

(プロジェクト研究「学力アセスメントの在り方に関する調査研究」)

学力アセスメントの動向と展望 ～ CBT 化に向けて ～

令和5年(2023年)3月

研究代表者 大金 伸光

国立教育政策研究所 教育課程研究センター長

はしがき

本報告書は、国立教育政策研究所令和4年度教育研究公開シンポジウム「学力アセスメントの動向と展望～CBT化に向けて～」の概要、講演録及び参加者アンケートをまとめたものです。

本シンポジウムは、国立教育政策研究所が実施しているプロジェクト研究「学力アセスメントの在り方に関する調査研究」（研究期間：令和3年度～令和5年度）の中間報告を兼ねています。

GIGA スクール構想の推進や CBT※による学力調査が国際的な標準となりつつある現状等を踏まえ、全国学力・学習状況調査等においても、CBT化に向けた検討・取組を進めることが急務となっています。こうした中、Society5.0に向けて、学習をめぐる新しい評価改善のサイクルを確立することが求められています。

本調査研究では、学習科学、教育測定、教科教育、データサイエンス及び実務的知見等を架橋して、CBT移行を展望した作問・結果分析の枠組みの在り方を検討するとともに、新たなPDCAサイクルの要となる学力アセスメントの改善充実にに向けた調査研究を推進しています。

シンポジウムは、オンライン配信による形式で行いました。760名を超える参加登録があり、当日は562名が参加しました。また、参加者アンケートから、たくさんの御感想と今後に向けた示唆を頂き感謝いたします。

本報告書によりその内容が更に多くの方々に広がり、教育に携わる全ての関係者に御活用いただけることを期待するとともに、更に本調査研究を深めて参りたいと考えます。

引き続き、皆様方からの厚い御協力、御支援を賜りますようお願い申し上げます。

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

令和5年(2023年)3月
研究代表者：大金伸光
教育課程研究センター長

目次

第1章 シンポジウムの概要	1
第2章 シンポジウム講演録	7
開会挨拶	
永山 裕二（国立教育政策研究所長）	8
趣旨説明	
大金 伸光（国立教育政策研究所教育課程研究センター長）	9
第一部	
基調講演	
学力アセスメントの展望と課題 — 全国学力・学習状況調査の CBT 化を契機に	
大津 起夫（大学入試センター 参与・名誉教授、国立教育政策研究所フェロー）	11
講演	
公的な学力アセスメントをめぐる国際動向について	
①PISA の進化について：特に CBT 化の観点から	
池田 京（OECD 教育スキル局シニアアナリスト）	17
②米国の公的な学力アセスメントについて	
石井 英真（京都大学大学院教育学研究科准教授、国立教育政策研究所フェロー）	24
第二部	
パネルディスカッション	
これからの学力アセスメント— CBT 化に向けて	31
コーディネーター	
白水 始（国立教育政策研究所初等中等教育研究部副部長・総括研究官）	
パネリスト	
袈岩 晶（国立教育政策研究所教育データサイエンスセンター総括研究官）	
渡辺 洋平（埼玉県教育局市町村支援部義務教育指導課長）	
佐藤 寿仁（国立大学法人岩手大学教育学部准教授、国立教育政策研究所 CBT スーパーバイザー）	
池田 京（OECD 教育スキル局シニアアナリスト）	
閉会挨拶	
大野 彰子（国立教育政策研究所 教育データサイエンスセンター長）	48
第3章 アンケート結果から得られた示唆	49

第1章 シンポジウムの概要

以下では、本シンポジウムの概要、参加者数、講演内容の要約を記す。

シンポジウム名

学力アセスメントの動向と展望 ～ CBT 化に向けて ～

開催日時

令和 4 年 11 月 30 日（水）15:00～17:00、オンライン開催（Zoom ウェビナー）

目的

GIGA スクール構想や CBT による学力調査が国際的な標準となりつつある中で、プロジェクト研究「学力アセスメントの在り方に関する調査研究」で得られた知見を踏まえながら、国内外の公的な学力調査の CBT 化をめぐる動向を中心に、学力アセスメントの発展可能性について考える機会とする。

プログラム

開会挨拶（15:00～15:05）

永山 裕二（国立教育政策研究所長）

趣旨説明（15:05～15:10）

大金 伸光（国立教育政策研究所 教育課程研究センター長）

基調講演（15:10～15:30）

「学力アセスメントの展望と課題ー全国学力・学習状況調査の CBT 化を契機に」

大津 起夫（大学入試センター 参与・名誉教授、国立教育政策研究所フェロー）

講演（15:30～16:00）

「公的な学力アセスメントをめぐる国際動向について」

①PISA の進化について：特に CBT 化の観点から

池田 京（OECD 教育スキル局シニアアナリスト）

②米国の公的な学力アセスメントについて

石井 英真（京都大学大学院教育学研究科准教授、国立教育政策研究所フェロー）

※事前録画

パネルディスカッション（16：05～16：55）

「これからの学力アセスメントーCBT化に向けて」

コーディネーター

白水 始（国立教育政策研究所 初等中等教育研究部副部長・総括研究官）

パネリスト

巖岩 晶（国立教育政策研究所 教育データサイエンスセンター総括研究官）

渡辺 洋平（埼玉県教育局市町村支援部義務教育指導課長）

佐藤 寿仁（国立大学法人岩手大学教育学部准教授、国立教育政策研究所 CBT スーパーバイザー）

池田 京（OECD 教育スキル局シニアアナリスト）

閉会挨拶（16：55～17：00）

大野 彰子（国立教育政策研究所 教育データサイエンスセンター長）

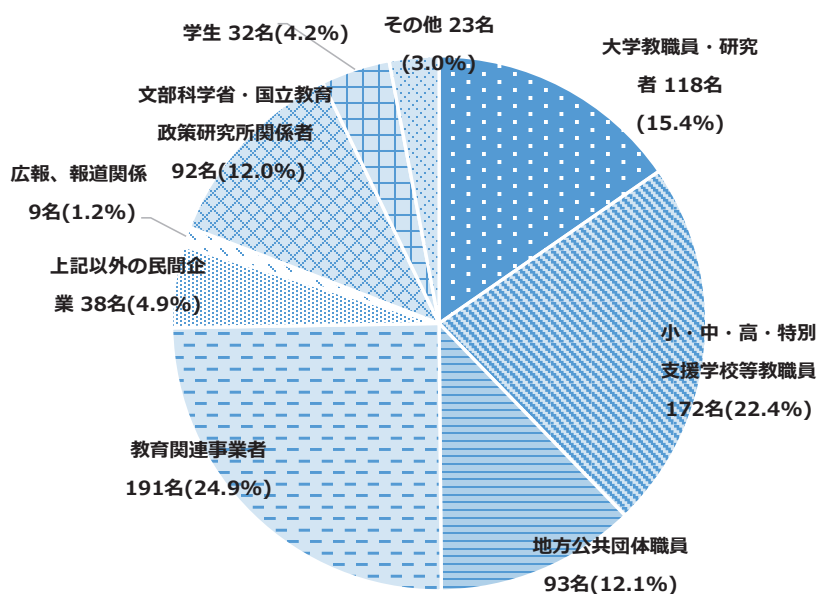
総合司会

岩間 裕美（国立教育政策研究所 教育データサイエンスセンター コンピュータ使用型調査推進課長）

参加者数

当日の参加者数は562名であり、参加申込者768名に対して、約73.2%の参加率であった。参加申込者の属性を図1に示す。

図1 参加申込者の属性



属性として最も多かったのは、「教育関連事業者」の191名(24.9%)、次いで、「小・中・高・特別支援学校等教職員」の172名(22.4%)、「大学教職員・研究者」の118名(15.4%)であった。教育産業界の関心が高くなっているとともに、教育の実践者や研究者、行政職員にも注目されていることが伺える。

講演内容の要約

◇基調講演

大津氏からは、文部科学省の「全国的な学力調査に関する専門家会議」の下に設けられた「全国的な学力調査のCBT化検討ワーキンググループ」の「最終まとめ」(令和3年7月16日)について、最終まとめに至るまでの経緯、内容、そして実際にどのようにCBT化を進めるかといったことについてお話があった。また、上記に関連して、問題配布の効率化、採点や集計の効率化・即時化、多様な出題形式への対応等のCBTの利点、情報インフラ整備のコストや機器操作能力の影響、同時実施の困難等のCBTの課題について提示された。また、分散実施の必要性を踏まえた調査設計や項目反応理論(IRT)について御紹介があった。最後に、データを活用するためには、結果が伝えていることを真摯に受け止めることが必要であり、調査の質の高さが求められるとの御意見を頂いた。

◇講演①

池田氏からは、PISAのCBT化について御紹介があった。まず、PISAの歴史について、2000年から3年ごとに実施されており、2006年からCBT(PISAではCBAと表現)が段階的に取り入れられ、2015年からCBTが主流となった経緯についてお話をされた。さらに、CBTのメリットとして、①「実生活」の文脈で学生のコンピテンシーを測定することができること、②革新的な新しいタイプの問題の導入、③テスト内容を学生の能力に合わせる適応型テストの導入、④反応時間やプロセスデータなどの付加情報を収集できること、⑤テスト問題の安全性、データの品質保証、調査実施に必要なリソースの削減といった観点を提示し、それぞれ具体的な分析や図表を示して説明された。一方で、課題点として、①筆記型とコンピュータ使用型の違いによる難易度の変化、②学校及び国間のコンピュータ性能の違い、③生徒のコンピュータ習熟度のばらつき、④適応型テストの導入による非ランダム化、⑤公開問題共有の難しさ、⑥進化していく調査への適応が挙げられ、各課題について、PISAがどのような対応を取っているか紹介された。

◇講演②

石井氏からは、米国の公的な学力アセスメントの経緯等を踏まえた上で、CBT化に向けた前提や論点についてお話し頂いた。まず、米国について、学力テストが注目を集めるのは1980年以降であり、特に各州における統一テストを軸としたアカウントビリティ政策の展開が中心であったこ

と、全米の学力向上政策の展開について説明された。また、標準テストなどの競争主義的な教育改革、及びパフォーマンス評価といったそのオルタナティブの模索について紹介され、全米学力調査（NAEP）について、どのような目的の下で行われてきたか、その変遷について説明された。また、日本と米国の状況を比較し、米国は競争的、結果責任を問うようなアウトプット機能が強い形であるのに対して、日本は学習指導の学力観や趣旨などを、調査問題を通じて伝達していくインプット、理念提示機能の性格が強いことが示された。最後に、CBT 化に向けた論点提示として、今の全国学力・学習状況調査の状況がどのように変化していくのかを踏まえた議論が必要であること、調査目的を検討しながら進めていく必要があることが論じられた。

◇パネルディスカッション

パネルディスカッションでは、前半の基調講演及び講演を振り返るとともに、それぞれの取組や実践のPDCA サイクルを充実させていくためには、どのようにあるべきかというテーマのもと、各パネリストの発表、及び議論が行われた。コーディネーターの白水氏からは、最初に全国学力・学習状況調査による教育施策の評価といった国の貢献可能性、指導改善といった学校での活用可能性、施策評価と指導改善の結びつけといった自治体での活用可能性の観点が提示された。

まず、渡辺氏より、埼玉県学力・学習状況調査の紹介、及び CBT の可能性についてお話があった。埼玉県学力・学習状況調査について、その特長として、児童生徒一人一人の学力の伸びを測定できることが挙げられた。また、埼玉県学力・学習状況調査を活用した学力向上施策について紹介された。最後に、上記調査の CBT 化について、メリットとしては、より精緻なデータの取得が可能になること、自動採点等の取組によって、短い時間で児童生徒にフィードバックが可能になることが提示された。渡辺氏の発表について、池田氏より、学力調査の結果を他の公的データと結合させることの有効性といった視点が提案された。

次いで、佐藤氏より、これまでの全国学力・学習状況調査を踏まえながら、現場の指導改善に近いところから CBT に対する期待についてお話を頂いた。まず、全国学力・学習状況調査が平成 19 年度から悉皆調査として実施され、平成 31（令和元）年度からは、新学習指導要領の考え方への各教育委員会や各学校の理解を促すため、主として「知識」に関する問題と、主として「活用」に関する問題に区分するといった整理を見直して、一体的に調査問題を構成するようになったことが紹介された。また、その一例として、中学校数学において、問題発見・解決の過程を重視した問題作成へと枠組みを変更したことが紹介された。また、実際に出題された小学校算数の問題をもとに、児童生徒の問題解決に関連するデータが得られていること、それによって教師の勤とデータが結びつくことで、指導改善に生かすエビデンスとなっているのではないかというお話を頂いた。最後に、CBT による全国調査を行うに当たって、現在の学力調査に係るサイクルから調査実施や結果分析の一連のプロセスが短縮されるのか、紙（PBT）では問えなかったことが CBT にすることで問えるようになるか、教科教育の先生方や専門家と一緒に検討しなければならないので

はないか、用いる統計的手法のメリット・デメリットの確認も含め、CBT化に向けて様々な専門家のもとでブラッシュアップする必要があるのではないかという論点を提示いただいた。

これらの発表を踏まえ、国の貢献可能性として、白水氏及び袈岩氏より、それぞれ班長を務めるプロジェクト研究「学力アセスメントの在り方に関する調査研究」の作問・結果分析班、測定技術班の取組を紹介した。白水氏からは、作問・結果分析班では、まず作問・結果分析のPDCAのサイクルを充実していくことを目指し、児童生徒の解き方のプロセスに焦点を当てたというお話があった。また、正答数別類型割合グラフを提示し、正答数が多い／少ない児童が、当該問題に対してどのような解答をしたのかという情報が得られることで、問題ごとに学力層の解答プロセスの思考の特徴が見える可能性が説明された。袈岩氏からは、測定技術班では、項目反応理論を全国学力・学習状況調査にどのように導入するかを検討しているとの説明があった。項目反応理論の特徴として、問題セットが異なっても能力を比較可能にすることが挙げられるとともに、分散実施に向けてどのようなことを行うのかが説明された。また、項目反応理論の概要について紹介された。

最後に、これまでの発表内容を踏まえて、池田氏、佐藤氏、渡辺氏の順にコメントを頂いた。池田氏からは、各問題がどのような能力を測定しているのかを、学校教員に向けて「見える化」することで、日々の実践に活用できるのではないかと、そしてそのような情報提供はコンピュータ化によって容易になるのではないかと御提案を頂いた。佐藤氏からは、児童生徒の特性をどのように測っていくのかと言うことに焦点をあてていくことに期待し、能力という言葉の意味について今一度説明することの必要性について御教示いただいた。渡辺氏からは、データをどのように使用するのか、どのように使うと効果的なのかといった点について、国研には好事例を示すなど道しるべとなってほしいという御意見を頂いた。

当日の発表資料や動画は、以下のウェブサイトより公開しているので、是非御覧いただきたい。

https://www.nier.go.jp/06_jigyousymposium/sympo_r04_01/

第2章 シンポジウム講演録

以下、シンポジウムの次第に従って、講演録を記載する。なお、講演録は追記修正を行ったため、シンポジウム当日の発言のままではないことを了承いただきたい。講演に伴う配布資料及び動画については、国立教育政策研究所のウェブサイト¹に掲載されている。

開会挨拶及び趣旨説明

国立教育政策研究所長 永山 裕二

国立教育政策研究所 教育課程研究センター長 大金 伸光

永山：皆さんこんにちは。文部科学省国立教育政策研究所所長の永山と申します。令和4年度教育研究公開シンポジウムの開催に当たり、一言御挨拶を申し上げます。

現下の状況から、今回もオンラインでの開催とさせていただきました。まず、皆様、大変お忙しい中、大勢の方々が本シンポジウムに御参加くださることに心よりお礼申し上げます。どうもありがとうございます。

国立教育政策研究所では、「学力アセスメントの在り方に関する調査研究プロジェクト」を令和3年度から来年度までの3年間の計画で、当研究所の教育課程研究センター及び教育データサイエンスセンターを中心に進めています。学校教育は、「GIGA スクール構想」による1人1台端末と高速大容量の通信ネットワークの整備が進む中、大きく変貌しつつあります。

また、PISA、TIMSS という国際学力調査は紙による調査ではなく、コンピュータを使った調査、CBT が標準となりつつあります。こうした状況などを踏まえまして、我が国においても毎年、小学6年生、中学3年生を対象に実施している、全国学力・学習状況調査の CBT 化に向けた取組が進められてきています。

さらに、政府の方針として、教育のデジタル化や教育におけるデータの効果的な活用、政策立案への反映などの重要性が強調され、文部科学省でも教育データの利活用による学び、指導支援の充実、EBPM の推進などに取り組んでいます。「学力アセスメントの在り方に関する調査研究」は、正にこうした動向に対応したものです。

今回のシンポジウムのテーマは、「学力アセスメントの動向と展望 ～CBT 化に向けて～」とし、まして、まず基調講演として、大学入試センターの大津様より「全国学力・学習状況調査の CBT 化検討ワーキンググループ 最終まとめ」の紹介も含め、「全国的な学力アセスメントの展望と課題」について御講演を頂きます。

次に、公的学力アセスメントをめぐる国際動向につきまして、OECD の池田様より「OECD 生徒の学習到達度調査 PISA の進化、特に CBT 化の観点から」というテーマで御講演を頂きます。また、京都大学の石井様より、アメリカの公的学力アセスメントと日本の学力調査の CBT 化に向けた論点提示を頂きます。

そして、締めくくりとして、これからの学力アセスメントについて学力調査に関わる様々な立場の方に御参加いただきまして、パネルディスカッションを予定しています。

全体で2時間ほどのプログラムになります。

本日のシンポジウムが皆様にとりまして、学力アセスメントに関する論点や、今後の展望を整

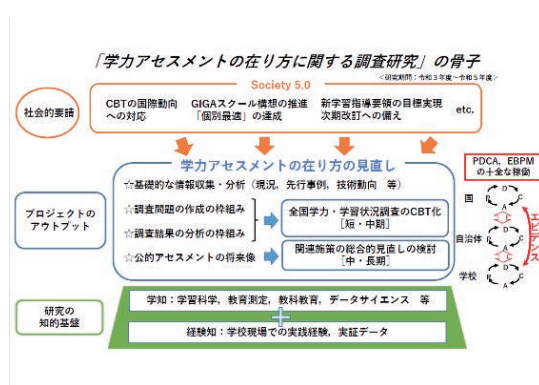
¹ https://www.nier.go.jp/06_jigyuu/symposium/sympo_r04_01/

理、共有する上で有益な機会となり、これからの日本の教育の質の向上に生かされますことを切に願っております。

最後になりましたが、本研究所への引き続きの御支援をお願いいたしますとともに、皆様方の一層の御活躍を祈念し、開会に当たっての御挨拶といたします。

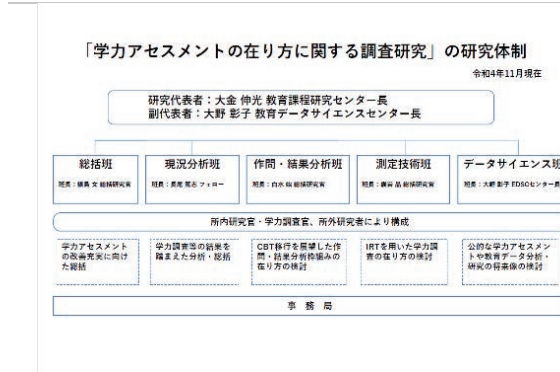
御清聴どうもありがとうございました。どうぞよろしくをお願いいたします。

大金：皆様、こんにちは。国立教育政策研究所 教育課程研究センター長の大金でございます。本日のシンポジウムの趣旨説明をさせていただきます。



本日のシンポジウムは、国立教育政策研究所が実施しております「学力アセスメントの在り方に関する調査研究」プロジェクトの中間報告を兼ねています。GIGA スクール構想の推進や、コンピュータによる調査、いわゆる 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・利点を生かした出題や、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 EBPM のさらなる推進を図るため、全国学力・学習状況調査の CBT 化に向けた検討・取組を

進めることが急務となっております。こうした中、Society5.0 に向けて、学習をめぐる新しい評価改善のサイクルを確立することが求められていると考えております。このため、本プロジェクトでは、学習科学、教育測定、教科教育、データサイエンス及び実務的知見等を掛け合わせて、CBT 移行を展望した作問・結果分析の枠組みの在り方を検討するとともに、新たな PDCA サイクルの要となる学力アセスメントの改善充実に向けた調査研究を推進しています。



プロジェクトの研究体制としましては、総括班、現況分析班、作問・結果分析班、測定技術班、データサイエンス班の5班からなり、所内研究官、学力調査官、所外研究者により構成しております。

シンポジウムの趣旨

【シンポジウムの趣旨】

GIGAスクール構想やCBTによる学力調査が国際的な標準となりつつある中で、プロジェクト研究「学力アセスメントの在り方に関する調査研究」で得られた知見を踏まえながら、国内外の公的な学力調査のCBT化をめぐる動向を中心に、学力アセスメントの発展の可能性について考える機会とする。

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

今回のシンポジウムのテーマは、「学力アセスメントの動向と展望 ～CBT 化に向けて～」であります。毎年、文部科学省で実施している「全国学力・学習状況調査」が、今後、段階的に紙による調査からコンピュータによる調査、CBT による調査に移行していく中で、日本より先行して CBT で行われている国外の学力調査の事例やプロジェクト研究の経過報告、パネルディスカッションを通して、公的な学力アセスメントの在り方や発展の可能性を考える機会としたいと思います。

第一部 15:10～16:00

◇基調講演◇

学力アセスメントの展望と課題 全国学力・学習状況調査のCBT化を契機に

◇講演◇

公的な学力アセスメントをめぐる 国際動向について

- ① PISAの進化について：特にCBT化の観点から
- ② 米国の公的な学力アセスメントについて

から御講演を頂きます。OECDが進めているPISA・生徒の学習到達度に関する調査におけるCBT化への歩みや、そのメリットや課題をOECD教育スキル局池田京シニアアナリストから御紹介を頂き、次に、米国の公的な学力アセスメントの変遷等について、京都大学大学院教育学研究科石井英真准教授から、こちらは事前収録による講演になります。

第二部 16:05～17:00

◇パネルディスカッション◇

これからの学力アセスメント ーCBT化に向けて

全体を二部構成とし、第一部は、国内外の公的な学力調査のCBT化をめぐる動向を中心に、次の3つの講演を行います。まず、文部科学省の「全国的な学力調査のCBT化検討ワーキンググループ」の最終まとめの内容や議論の紹介、さらに、CBT化を契機とした全国的な学力アセスメントへの展望と課題、期待などについて、CBT化検討ワーキングの主査でいらっしゃる、大学入試センター参与・名誉教授の大津起夫先生から基調講演を頂きます。次に、公的な学力アセスメントの国際動向について、お二人

第二部「パネルディスカッション」では、「これからの学力アセスメントーCBT化に向けて」と題して、様々な視点から御議論を頂きます。地方自治体の観点から、埼玉県教育局の渡辺洋平義務教育指導課長、教科教育の観点から、岩手大学教育学部の佐藤寿仁准教授、教育測定の観点から当研究所の巖岩晶総括研究官、国際的な視点からOECDの池田シニアアナリストに御参画を頂きます。

それではどうぞよろしくお願い申し上げます。

第一部

基調講演 学力アセスメントの展望と課題

—全国学力・学習状況調査の CBT 化を契機に

大津 起夫：大学入試センター 参与・名誉教授、国立教育政策研究所フェロー

学力アセスメントの展望と課題

—全国学力・学習状況調査の CBT 化を契機に

大津 起夫
大学入試センター

国立教育政策研究所 令和4年度教育研究公開シンポジウム
学力アセスメントの動向と展望 ～CBT化に向けて～
2022年11月30日 15:00～17:00 (オンライン) 事後修正 2022年12月5日

大津：よろしくお願いします。

御紹介にあずかりました、大学入試センターの大津です。本日は学力アセスメントの展望と課題というタイトルで講演させていただきます。

皆さん御存じかと思いますが、CBTという言葉が出てきますが、これはコンピュータ ベーストテイスティングでという意味で、通常はコンピュータ端末でする試

験のことを表しています。

全国学力・学習状況調査、以後、全国学力調査と省略させてもらいますけれども、それを次第に CBT 化していくということで、そのためにどういうことを考えなければならないかということに関して報告をまとめました。

概要

- 「全国的な学力調査のCBT化検討WG 最終まとめ」(令和3年7月16日)について
「全国的な学力調査に関する専門家会議」の下に設けられたワーキンググループの報告
- CBTと教育測定論
- データを活かすためのマインド

全国的な学力調査の CBT 化検討ワーキングというのは、専門家会議の下に設けられたワーキンググループです。昨年の夏、7月に最終まとめというものを公表しました。今日は主にその内容について報告します。詳細に関しては、報告書が文部科学省の方のウェブページに出ていますので、見ていただければと思います。それからいろいろ概念的な新しい内容が従来

の学力調査と違う側面というのを考えなきゃいけないということで、CBT と教育測定、要するに尺度をどういうふうに構成するかということに関して、これ以降の御発表でもいろいろ報告があると思うのですが、簡単に説明させていただきます。

それから、最後に時間があれば、そのデータを生かすためのマインドとか心構え、その技術的な面以前の問題としてどういうことが重要かということに関して、若干触れられればと思います。

CBT化検討WG 最終まとめ

- ・報告書（令和3年7月16日）は文科省のWebサイトにある

・「最終まとめ」までの経緯：

前年に「中間まとめ」を公表（令和2年8月28日）

① 調査の目的と実施方法は表裏一体の関係

② CBT化に向けた体制整備の必要性（テスト理論やデータ分析の専門家が国研にいたることが望ましい）→ 教育データサイエンスセンターの設置

③ 端末による学習環境への習熟、発達段階への考慮

④ CBT化による現場への負担の考慮

と実施方法は表裏一体の関係だとの指摘です。どういう調査目的をとるかということによって、どういう技術的工夫を凝らすかが関係しますよということです。2番目として、CBT化に向けた体制整備、特にテスト理論とかデータ分析の専門家が国研にいたることが望ましいということを指摘しました。これは、かなり体制が関係者の御尽力によって整備されて、教育データサイエンスセンター等が設置されました。3番目として、そのコンピュータ端末による学習環境への習熟とか、発達段階を考慮しないとイケませんねということです。それから4番目としては、CBT化にすると機械操作とかトラブルとかいろいろなことが発生すると思われまますので、その現場への負担が余り重くならないように考慮しないとイケないということを指摘しています。

CBT化検討WG 最終まとめ

- ・背景： GIGAスクール構想、PISA等のCBT化
令和3年7月時点での方針
- ・教育測定、教育社会学、教科指導、特別支援教育などの専門家により、専門的・技術的観点から検討
- ・本体調査（悉皆）と補完調査（経年変化分析調査：保護者調査を含む）を二本柱とし、役割分担
- ・悉皆調査：教育理念・目標の呈示、学習指導の改善、教育施策の検証・改善
- ・経年変化分析調査：全国的な学力状況の経年変化の把握
- ・各調査の目的に即して、最適な方法を設計しCBT化をする

それで CBT を実際にやるためのハードルがかなり低くなったということが前提条件としてあります。それから、先ほども御紹介にありました、PISA 等の国際学力調査がおしなべて CBT 化が進んでいる。それを考えると、当然日本においても、考える必要があるでしょうということです。それから、このワーキンググループですけれども、教育測定とか教育社会学、教科指導、それから特別支援教育などの専門家の先生に入っていて、その専門的技術的な観点からいろいろな検討をしていただきました。

重要なポイントなのですが、本体調査、これは小学校6年生、中学校3年生、原則としてみんなに参加してもらって調査ですね。それと補完調査、これは国がサンプリングして数百校の学校を選んでそこに対してする調査なのですが、これは経年変化分析調査というものと保護者調査という両方を含んでいるのですけれども、この本体調査と補完調査というのを二本柱としてそれぞれに役割分担させるということが必要でしょう。悉皆調査に関しては、これは教育理念とか目標を、その課題ですね、問題・設問を通じて提示するし、その結果、その設問の内容等を利用して学習指導の改善を行う。それから、特に地方のという意味合いが強いかと思うのですが、教育施策の検証、それから改善に役立てるといって、そういう役割があるでしょう。一方、補完調査のうちの経年変化分析調査ですけれども、こちらはですね、全国的な学力状況がどういうふうに変化しているか、それを精密に把握するということに重きを置くということです。少なくとも現状の悉皆調査は毎年、毎年問題が全部違いますから、作られる先生は余り極端に難しさが変わらない

CBT化検討ワーキングの最終まとめは、去年7月に出したのですけれども、検討はそれ以前から始めていて、実際にその中間まとめの論点整理というのを、その前の年、令和2年の8月に公表しています。令和2年の前期にかなり集中的に議論をしました。令和2年の段階で、次の4点を特にその論点として提示しているわけです。

1番目として当然なのですが、調査の目的

その CBT 化検討ワーキンググループの最終まとめの内容なのですけれども、ここに書いているようなことが特に指摘されています。背景としては、先ほどの所長の御説明にもありましたけれども、まず GIGA スクール構想がかなり進展しました。学校の中のコンピュータ、端末整備も進みましたし、その学校間をインターネットにつなぐネットワークも大幅に強化

ように配慮されていると思うのですが、直接昨年と今年とどちらが良かったか、設問正答率は比較できますけれど、それが難しさとしてどうなのかというのは、直接には少し比較しづらいわけですね。経年変化分析調査の方はいろいろ出題方式に工夫を凝らして、精密な比較が可能になるように工夫してありまして、その仕掛けを CBT に持っていくということです。それぞれの調査の目的に即して最適な方法を設計して、CBT 化をするという趣旨の報告をしています。

CBT化検討WG 最終まとめ（ポイント）

- 令和3年度以降、試行・検証を繰り返し、段階的にCBT化
- 令和6年度から、全国学力・学習状況調査に順次CBTを導入
- CBTへの移行、運用のための国立教育政策研究所の体制強化
- 学校・教職員のための技術サポート
- ネットワーク環境の整備

あとは、どういうふう実際に CBT 化を進めようかということなのですが、令和3年度、昨年ですけれども、それ以降、試行・検証を繰り返して段階的に CBT 化を行うということです。試行・検証は既に令和3年度以降ですね、少しずつ文科省の方で行われています。今年度もそれから来年度も、来年度は結構大規模にやると思いますが、実際、既に幾つか

仕事が行われています。

それから、本番への CBT の導入ですけれども、令和6年度から全国学力・学習状況調査本番に順次 CBT を導入していくと。それから増やしてはどうかということを提言しているわけです。それから、CBT への移行、運用のために国立教育政策研究所の体制を強化する。これは所長も言及されていたその体制が次第に整備されています。それから、先ほどもありましたけれども、もし実際に CBT をみんなやるということになれば、先生方にただやってくださいというわけにもいかないで、技術的なサポートは当然必要だよということですね。

それから、最後のところは GIGA スクールでかなり改善されましたけれども、ネットをつながない状態で CBT をやるというのができないわけではないですけれども、非常に困難だし、手間がかかります。やはり一つずつのコンピュータはネットワークにつないで問題の配信とか改修を行った方が望ましい、そうせざるを得ないであろうということで、ネットワーク環境の整備も絶対必要だよということを書きました。

CBT化検討WG 最終まとめ（ポイント）

- 悉皆調査：
 - 全国同日一斉実施 → 一定期間内の分散実施へ
 - 各回一種類の冊子 → 比較可能な問題セットを複数用意
 - 児童生徒質問紙調査は、令和6年度を目途にオンライン回答方式を全面導入
 - 教科調査については、中学校から先行し令和7年度以後できるだけ速やかに導入
- 経年変化分析調査：（3年ごとに実施）
 - 現行の調査設計を基本とし、令和6年度からCBTを導入（紙筆テストを経過的に併用）
 - さらに専門的・技術的な検討を期待
- 工程：
 - 目標時期や手順等は、技術発展の状況や試行・検証結果を踏まえて変更もある。
 - （予想外の困難が生じたり、技術環境の変化によってあらかじめ考えていたことが陳腐化するかもしれない）

ポイントですけれども、悉皆調査については、現在の悉皆の調査というのは全国同日に一斉実施です。これをただそのままです、100万プラス100万ですから、それを同時に機械を用いて、コンピュータを用いてやるというのは、少なくとも現状で考える限りかなり困難であると。それなので、一定期間、時間の幅を設けて、その期間内に分散実施を

するようにしたい。現状では毎回異なる問題を出しているのですが、要するに一回の試験ですね、1年に行われる試験ってそれぞれ中学校小学校それぞれに一つの科目について1種類の問題ですね、一つの問題、同じ問題冊子を全員解いてもらっているのですが、分散実施すると、全員同じとなるとやはりタイムラグがあるので、それは少し問題があるのではないかとということですね。

学力を比較可能な形に、何か工夫した問題セットの冊子をですね、それぞれ複数用意して、例えば一番典型的に考えられるパターンというのは、実施日によって問題を変えるということで

すけれども、そういうふうな複数の問題セットを使う体制を用意する必要があるねということ
を指摘しています。

それから、その学力調査と一緒に行われている児童生徒質問紙、アンケートですね、生徒さん
にするアンケートに関しては、これはですね、令和6年度を目途にオンライン回答方式を全面導
入するという事です。教科調査は令和6年度に全部というわけには少しいきそうもないので、
機器操作の習熟度の問題もあるので、まず中学校から先行して令和7年度以降、できるだけ速や
かに導入するというプランを提示しています。もう一つですね、経年変化分析調査、毎年やっ
ているわけではなくて、3年ごとに実施しているのですが、これは既にIRTでコントロールされ
た複数の分冊方式をとっています。現行の調査設計を基本として、令和6年度からCBTを導入す
る。ただ、最初は現行と同じようなやり方、紙のテストを経過的に併用して、CBTと紙で性質が
違ったりしないかどうかということを確認しながら、順次CBTに移行するという事です。さら
に、専門的技術的な検討を期待すると書いてあります。

工程ですけれども、これ、一般的なコメントではあるのですが、目標時期ですね、何年にどこ
までやるとか、それからどの段階でどういうふうに進めるとかというのは、手順ですね、技術発
展の状況とか試行・検証結果を踏まえていろいろ予期しない困難とか発生する可能性もあるの
で、変更もあり得ますよということですね。

特に実際に大量に機械を使った調査をやってみると、予想外の困難が生じたりするかもしれ
ませんし、特に数が膨大でするので、何が起るかわからないという側面があります。それから、
技術環境の変化によっては、あらかじめ考えていたプランというのが実は陳腐化してしまうと
いうことも、発展が速いですから、そういうこともあるかもしれません。

CBTと教育測定論

- CBT化の利点
 - 問題配布の効率化
 - 個別対応も頑張れば可能（学力調査より学習支援において特に有益だろうか）
 - 採点・集計の効率化、即時化（記述式採点の自動化は研究途上）
 - 多様な出題形式、特別な配慮が必要な児童生徒への対応
- CBT化の問題点
 - 情報インフラ整備のコスト（GIGAスクール構想でCBTの実現化が現実的になったが、今後の情報環境維持にもコストはかかる）
 - 学力以外の機器操作能力の影響
 - 大規模な調査で、機器を支障なく稼働させることは大変
100万人+100万人規模での同時実施は、かなり困難

それで2番目の話題に入るのですが、
CBTと教育測定論ということ。先ほ
ど、複数の分冊方式をとるとい
うようなことも言いましたが、
それは次のスライドで触れま
す。ここでまず、では、何で
CBT化なんてやるのだという話
ですけれども、現状100万人に
問題を配布する、それが相当大
変なロジスティクスを必要とす
るわけです。もしCBTが導入さ

ればですね、それもそれで大変は大変なのですが、電子的に問題を全部配信することができます。それから、個別対応、これは実際やるという計画が具体的に今挙がっているわけではありま
せんけれども、一人一人の学習履歴とか解答状況ですね、それに応じて違う問題を出すというよ
うなこともがんばれば可能です。ただ、これは具体的に日程に載っているわけではありませ
んで、むしろ学力調査というより、個別の学力支援において有益なのではないかなとは思
います。それから、当然のことですけれども、採点・集計が効率化できますし、即時化、少なくとも多肢
選択とか簡単な解答に関しては即時化できるだろうと。ただし、一般的な記述式の採点とい
うのは、なおですね、研究途上です。これはもう少し研究の成熟を待たないと業務として運営できる
レベルにはならないと思います。それから、動画とか音声とか多様な出題形式ができます。重要
な側面は、特に障害とかがあつて、特別な配慮が必要な生徒さんたちにそのハンディキャップに
応じた柔軟な出題形式をとるといことも可能かと思ひます。これもまた、いろいろ研究途上の
問題面だと思いますけれど、そういう工夫も可能であるということですね。

ただ、いいことばかりでなくて、当然問題もあります。一つ重要なことは情報インフラ整備に
お金がかかります。

GIGA スクール構想でかなり環境は良くなりましたけれども、少なくとも CBT をやる限り、ずっとその環境は維持しないとイケません。当然、機械は壊れますから、その情報環境維持にもお金、コストはかかるということになります。それから、先ほど少し触れましたけれども、実際にコンピュータを使って解答してもらった場合に、学力以外の機器操作の能力を実は測ったりしていないか、そういう心配はあります。だから、そこが問題ないように、十分に機器操作に習熟する、生徒さんが習熟するということが重要だということです。それから、これ何遍も言っていますけれど、大規模の調査で機器を支障なく稼働させるということは相当大変です。

特に小学校 100 万人、中学校 100 万人の規模での同時実施ですが、一斉にやるというのはまず無理だと思います。

CBTと教育測定論

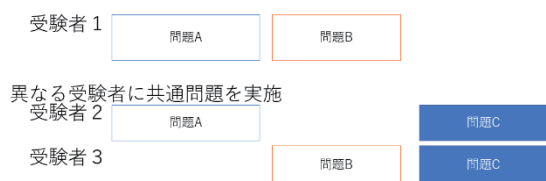
- 報告書の提案：
同日一斉実施から分散実施へ → 異なる問題セットの利用
「教科学力」ベースでの比較
「この問題ができた、あの問題ができなかった」 →
設問の正誤パターンから教科学力を推定し、その尺度で比較
異なる問題への解答結果を比較するためには、ある種の抽象化が必要
- IRT（項目反応理論）等の「テスト理論」とよばれる手法は、このようなテスト設計のための道具（1970年代にIRTの大枠は作られた）

少し抽象的な概念の尺度を設定して、その尺度を推定すると。その尺度の上で違った問題に解答した生徒さんたちの学力を比較するという考えをとらざるを得ません。IRT とか項目反応理論と言われますけれど、この後も説明あると思いますが、そういうテスト理論と呼ばれる手法というのは、このようなテスト設計のための道具として整備されたものです。

それで報告書の提案ではですね、その同日の一斉実施から、分散実施にする。そうすると異なる問題セットを利用して成績を比較するということが必要になると。それからこれはですね、この問題ができた、あの問題ができなかったという、即物的と言いますか、そういう具体的な比較から、もう少し抽象的な考え、問題の正誤パターンからですね、教科学力という

CBTと教育測定論

- 注意点：IRTの手法自体が異なるテストの比較を可能にするのではない。事前のテスト実施が難易度の評価のために必要。問題Aと問題Bを比較するには…
同じ受験者に異なる問題を実施



例えば、問題 A と B の難しさを比較するために一番シンプルな方法は、同じ受験者に問題 A と問題 B を両方解いてもらって、どちらがどれだけできたかというような情報をあらかじめ集めることです。もう少し凝ったやり方としては、受験者 2 には問題 A と問題 C を出して解いてもらう。それから受験者 3 には問題 B と問題 C を解いてもらう。そうすると、問題 C は共通に入っているので問題 A と問題 C を比較。それから問題 B と問題 C を比較することによって、間接的に問題 A と問題 B を比較するというようなことが可能になります。

要するに、事前にそういう調査をしておかないとイケないということです。

注意点なのですが、IRT はしばしば言及されますけれども、その手法というのは情報を整理するための方法で、その手法自体が異なるテストの比較を可能にするというのではないです。重要なのは、事前にテストを実施しておいて、難易度の評価のために、試験問題の難易度とかその特性をあらかじめ調べておく必要があるということです。

データを活かすためのマインド

- 「調査結果（数字）によるダメ出し」を受け入れられるか？
- もっときつく言うと、「熱意や努力が結果の良さを保障しない」ことを受け入れられるか？
- 調査には（必ずしも好意的ではない）関係者を説得するだけの質の良さが必要。調査の質にふさわしい受け止めが重要
- EBPM（Evidence Based Policy Making）のための条件

を払えば払うほど、その結果に対して否定的な答えが出たときに、そう素直には受け取れないことが多いのではないかなと思います。

逆に言えばですね、調査には必ずしも好意的ではない関係者を説得する質の良さが必要かなと思います。エビデンス ベースド ポリシー メイキング（Evidence Based Policy Making）としばしば言われますけれど、それが成立するためにはそういう素直さが必要かなと思います。

データを活かすための前提なのですが、これは分析法とかどういう情報機器を使うかという以前にですね、その調査結果による結果としてですね、ダメ出しが出るわけです。それを受け入れられるかということなのですね。もっときつく言うと熱意とか努力が結果の良さを必ずしも保障しないということを当事者が受け入れられるかということです。特に熱心で十分な努力

データを活かすためのマインド

- 人間の直観の強み、弱み：データに対して直観的に正しい判断ができるのなら、統計的手法は不要だが…
- 直観的判断への批判 Clinical versus actuarial judgment (Dawes, Faust, Meehl, 1989). 「臨床的判断」（症例についての直観的判断）と「計量的判断」の比較
- 統計的手法の役割： Controlled Magical Thinking (Diaconis, 1985) Magical Thinking (魔術的思考) とは、断片的な情報をつなぎ合わせて、ストーリーを作り出す人間の荒々しい力。神話や科学理論や迷信を生み出す力。この力を飼い馴らすための統計的手法の役割

ほぼ時間がなくなっただと思いますので、最後はですね、その人間の直感ですね、データを扱う際に特有の弱みみたいなものがあります。それに関しては、過去にいろいろ既に研究が行われて、重要なと思う研究を2つ紹介しておきました。

資料は、最後の方にその文献は挙げておきました。関心のある方は御覧ください。以上です。

資料

- 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21日）
https://www.mext.go.jp/a_menu/shotou/gakuryoku-chousa/1344321.htm
- R.M.Dawes, D.Faust & P.E. Meehl (1989). Clinical versus actuarial judgement. *Science*, 243, 1668-1674.
- P. Diaconis (1985). Theories of data analysis: From magical thinking through classical statistics. in D. C. Hoaglin, F. Mosteller, & J.W. Tukey eds. *Exploring Data Tables, Trends, and Shapes*, 1-36, Wiley. 探索的データ解析をいかに行うべきか。専門家向けの助言。

①PISA の進化について：特に CBT 化の観点から

池田 京：OECD 教育スキル局シニアアナリスト



PISA
Programme for International Student Assessment
PISAの進化について：特にCBT化の観点から

池田 京
シニアアナリスト
OECD教育スキル局



池田：PISA の進化について、特にコンピュータ化の観点からお話したいと思います。

》 PISAコンピュータ使用型調査(CBA)：歴史

- PISA調査は、2015年調査で筆記型調査(PBA)からコンピュータ使用型調査(CBA)へと切り替わった。
- PISA2018年調査から多段階適応型テスト(MSAT)を導入。



PISA では、CBA、コンピュータ ベースド アセスメントという言葉を使っています。PISA 自体は 2000 年に始まり、3 年おきに調査は行われています。2022 年と 2018 年の間が 4 年空いているのは、パンデミックの影響で調査が 1 年遅れたため、ここだけ 4 年の開きがあります。

現在は 80 か国以上の国と地域が PISA に参加し、読解、数学、科学、そして革新

分野という毎回変わる分野、計 4 分野について 15 歳の生徒の調査を行っています。PISA 自体は 2000 年から始まりましたが、2006 年から徐々に CBA を取り入れてきました。

PISA の 2006 年、2009 年、2012 年ではオプションとして希望する国が、科学の CBA や、読解、問題解決能力、数学、読解の CBA を実施しました。そして、2015 年から PISA の主な調査方法がペーパーベースからコンピュータに変わりました。このときはほぼほとんどの国が、参加国がコンピュータに移り、10 か国ほどがまだコンピュータに移れないということでペーパーベースで行われました。

そして、2018 年からは多段階適応型テスト、アダプティブテストが導入されました。まず 2018 年には読解力に適応型テストが導入され、2022 年現在は数学と読解両方に導入され、次の 2025 年には科学にも適応型テストが導入される予定です。



PISAコンピュータ使用型調査
メリット

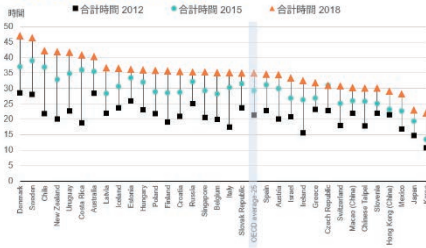
次に、コンピュータ使用型の調査のメリットを幾つか御紹介したいと思います。

メリット(1)

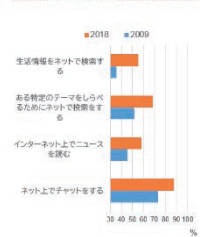
可能なこと

- 現在の「実生活」の文脈で生徒のコンピテンシーを測定することができる。

インターネットを利用した週あたりの時間数



以下の活動を週に数回以上行った生徒の割合(OECD平均値)



に移ってきました。そのことを示すデータがここにあります。2012年に比べて2018年、生徒のインターネット使用時間が大変伸びましたし、また生活情報をネットで検索するとか、インターネットでニュースを読むとか、ネットでチャットをするというような活動がどんどん増えてきています。

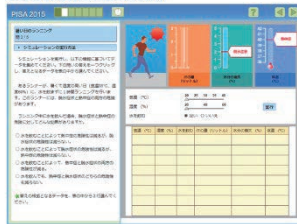
そういう意味でも、デジタル化が進む中で、コンピュータベースのアセスメントに移るということは、実生活の文脈で生徒のコンピテンシーを測るという意味で大事になってきます。

メリット(2)

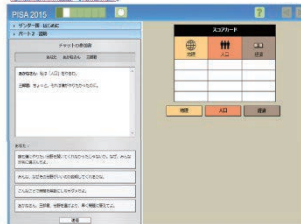
可能なこと

- 今日の「実生活」の文脈で学生のコンピテンシーを測定することができる。
- 革新的な新しい問題を導入する。

暑い日のランニング(リンク)



ザンダー国(リンク)



うのは、2015年の科学の問題です。画面の左側に問題がありまして、右側にシミュレーションをする場所があります。生徒はシミュレーションをすることによって問題に答える設定です。

こういったシミュレーションというのは、今までペーパーベースではできなかったタイプの問題です。2015年にコンピュータに移ることによって、こういったシミュレーション型の問題を導入することができました。右側の「ザンダー国」というのは、2015年に行われた共同問題解決能力、コラボレーティブ プロブレム ソルビング (Collaborative Problem Solving) の問題の1例です。これは生徒たちが他の生徒と協力して問題を解く能力というのを測っています。

他の生徒といったときに、これはコンピュータでジェネレート (generate) された他の二人と問題を解くということで、チャットベースで問題を解いていきます。生徒が答えると、それに対してあかねさんとかもう一人の男の子が答えて一緒に問題を解いていくという形になります。

こういった形の問題も、コンピュータに移行することによって可能になりました。

一番大きなものとしては、PISAはそもそも生徒のアプリケーション能力、どれだけ知っている知識や技能を実生活の中で使えるか、ということのを測っていますので、実際の実生活の文脈で生徒のコンピテンシーを測ることがとても大事になってきます。

ここ20年、デジタル化が進み、生徒の活動がペーパーベースからデジタルベース

次に、新しいタイプの問題を導入することができます。これは読解、数学、科学といったPISAのコアの分野でも言えることですし、また革新分野でも言えることです。

コンピュータベースでPISAを行っている国のみが革新分野に参加することができます。

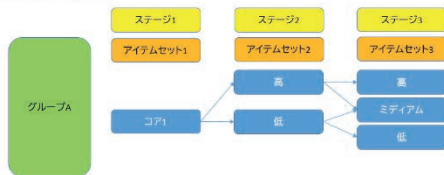
左側にある「暑い日のランニング」とい

メリット(3)

可能なこと

- 今日の「実生活」の文脈で学生のコンピテンシーを測定することができる。
- 革新的な新しい問題を導入する。
- テストの内容を学生の能力に合わせる - テストのターゲティングとエンゲージメントを高める。

多段階適応型テスト(MSAT)



ありまして、最初はコアと言われているアイテム（項目）セットを生徒に与えます。コアは普通のレベルの問題で、これらの問題の正答率によって2番目のステージでもっと難しい問題が与えられるか、若しくはもっと易い問題がもらえるかというのが決まってきます。

適応型テストの利点は、短い時間にその生徒の能力に合った問題に答えてもらうことで、よりテストの精密度が上がるということです。また、生徒も自分の能力に合った問題をもらうことによって、エンゲージメントが高まるということがわかっています。

そして、3番目のメリットとしては、適応型テストが導入できるということです。適応型テストというのは、生徒の能力に応じて、問題のグループ、どんな問題が与えられるかが変わってくるということです。

PISAのテストは2時間です。1時間受けた後に少し休みがあってもう1時間ということで、ここに示したのが最初の1時間様子です。その1時間を3つに分けて

メリット(4)

可能なこと

- 今日の「実生活」の文脈で学生のコンピテンシーを測定することができる。
- 革新的な新しい問題を導入する。
- テスト内容を学生の能力に合わせる - テストのターゲティングとエンゲージメントを高める。
- 反応時間やプロセスデータなどの付加情報を収集する。

ラバツイ鳥 (リンク)



OECD加盟国におけるナビゲーション行動グループの生徒の割合と平均得点との相関関係

群衆に最適化されたナビゲーション		積極的に活用するナビゲーション		限定的なナビゲーション		ナビゲーションなし	
Corr.	S.E.	Corr.	S.E.	Corr.	S.E.	Corr.	S.E.
0.69	(0.12)	0.43	(0.09)	-0.28	(0.09)	-0.32	(0.14)

リンク

て間違っただのかというのをプロセスデータから導き出すことができます。ここにある問題は、2018年の読解の例です。2018年の読解で新しくフォーカスしたのが、生徒がどうやってナビゲーションをするかという点です。紙媒体だと大抵読むのが最初のページから終わりのページまでリニア（直線的）ですけれども、コンピュータというのは生徒がいろいろなページをクリックし、そのプロセスを見ることができます。

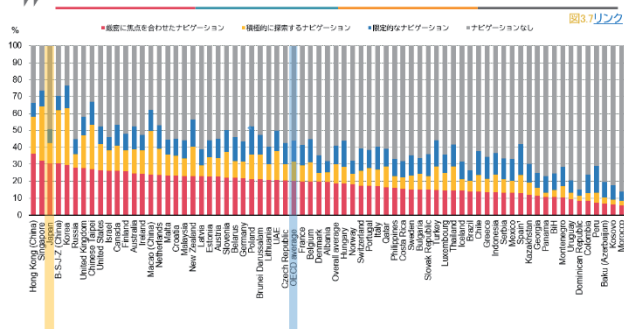
この問題もブログだとか、いろいろなページを見ることによって、生徒は問題を答えることが可能になります。どのようなナビゲーションを生徒がしたかというのをまず指標化し、4つのグループを作りました。ある一つの生徒のグループは、厳密に行くべきページのみを見た生徒たち、そして黄色の部分、積極的に検索するというのは、いろいろなページに行かなくても答えられる問題でも、いろいろなページを見た生徒たち。

3番目の限定的なナビゲーションのグループというのは、余りいろいろなページに行かなかった生徒たち。4番目のグループというのが全くナビゲーションがなかった生徒たち。ここに書いてあるコリレーション (Corr.) というのは、これが生徒の実際の成績とどう関係しているかというのを示してありまして、最初の厳密に焦点を合わせたナビゲーションをした生徒の割合が多いほど、国の平均点が高い傾向がみられました。

4番目に可能なこととして、紙では得られなかった追加情報を収集することができます。紙では、最後の正答、何が間違っただか、正しい答えかということのみがわかりますが、コンピュータではプロセスデータと言われる問題の答えにたどり着くまでのプロセスを見ることができます。

間違っただかというのをプロセスデータから導き出すことができます。

ラバ・ヌイ:ナビゲーション行動



これがこの4グループの割合を国際的に比較したものです。

この情報から先生たちは、どのようなナビゲーションを教えた方がいいのかというのが分かると思います。

メリット(5)

可能なこと

- 今日の「実生活」の文脈で学生のコンピテンシーを測定することができる。
- 革新的な新しい問題を導入する。
- テスト内容を学生の能力に合わせる - テストのターゲティングとエンゲージメントを高める。
- 反応時間やプロセスデータなどの付加情報を収集する。
- データやテスト問題の安全性を高める。
- データに関するあらゆるステップをフォローすることで、結果的にデータの品質を高めることができる。
- 調査実施に必要なリソースの削減(印刷、データ入力、コーディングなど)

そして、あと3つメリットを挙げると、まずはデータやテストの問題の安全性を高める、セキュリティですね。PISA というのは経年変化というのも見ているので、テスト問題がとても大事になってきます。

そのテスト問題のブリーチ (breach) がないかどうかというの、常に誰かがログインしたかっていうのもわかるので、安全性が高められます。また、データブリーチに

関しても誰がデータにアクセスしたのか、どの段階でデータが改ざんされたのかというのがわかるので、そういう意味では安全性は紙媒体よりも高まっていると言えます。

また、データに関するあらゆるステップを追跡することができるので、結果的にデータの品質を高めることができます。生徒のナビゲーションだけではなく、生徒がどれだけ問題に対して時間を費やしたかというのもプロセスデータから分かってきます。

ですので、生徒がほとんど時間を使わずに問題を20問ぐらいぱっぱと答えたりすると、それは生徒がエンゲージメントもなく適当に答えたということになります。前回2018年で1か国そういう問題が起こっているということがプロセスデータを見ることによってわかりました。プロセスデータを元に、どれだけ時間を生徒が1問に対して費やしているかという情報を見ることによって問題が判明しました。

そして、最後には調査実施に必要なリソースの削減。印刷とか、データ入力、コーディングなどのリソース削減ができると思います。



PISAコンピュータ使用型調査
課題

次に、幾つかの課題についてお話ししたいと思います。

課題(1)

- ・経年変化の測定 - 筆記型調査(PBA)とコンピュータ使用型調査(CBA)では項目の難易度が変わる可能性がある。
→スケーリングにおいて、一部の継続問題を別問題として扱う。

IRT スケーリングを通して調査しました。そしてもし問題がとても難しくなっていたり、とても簡単になっていたたりする場合は、見た目上は全く同じ問題でも違う問題、別問題としてスケーリングのときに扱われました。

特に数学の問題で数式を書く問題などは紙媒体だとそれほど難しくはないのですが、コンピュータを使うことによって、その数式を書き込むことが難しくなっていたりするので、数学の問題の幾つかは別問題として扱われました。

課題(2)

- ・経年変化の測定 - 筆記型調査(PBA)とコンピュータ使用型調査(CBA)では項目の難易度が変わる可能性がある。
→スケーリングにおいて、一部の継続問題を別問題として扱う。
- ・学校/国によるコンピュータ性能の違い
→学校のコンピュータのハードウェアとソフトウェアの設定が最低要件を満たしているかどうかをチェックするシステム診断アプリケーションの提供

参加する学校でそれを試してもらい、最低条件を満たしているということが確信された上で、テストを行うようにしました。また、幾つかの国では、ラップトップを購入して、学校に持っていく、そのラップトップを使って PISA に参加してもらおうということも行いました。

課題(3)

- ・経年変化の測定 - 筆記型調査(PBA)とコンピュータ使用型調査(CBA)では項目の難易度が変わる可能性がある。
→スケーリングにおいて、一部の継続問題を別問題として扱う。
- ・学校/国によるコンピュータ性能の違い
→学校のコンピュータのハードウェアとソフトウェアの設定が最低要件を満たしているかどうかをチェックするシステム診断アプリケーションの提供
- ・生徒のコンピュータに対する習熟度のばらつき
→練習のためにオリエンテーションを導入



とにもし独自の何か特別なコンピュータの機能、例えば先ほどのシミュレーションを使うものでしたら、そのシミュレーションができるためのオリエンテーションを各教科のテストの前に、追加しました。このオリエンテーションを通して、できるだけコンピュータの習熟度のばらつきを修正するという目的があります。

PISA は、経年変化の測定をしていまして、トレンドというのを測るのがとても大事です。

特に 2015 年に CBA に移ったときにはそれまで紙媒体で生徒が答えていた問題がコンピュータになったときに、果たして同じ難しさかどうかというのが大きな問題でした。ですので、予備調査の時点でその問題の難しさが紙媒体とコンピュータ使用のときとで同じかどうかというのを、

また、学校又は国間によるコンピュータ性能の違いというのが大きな問題になってきます。

学校のコンピュータのハードウェアとソフトウェアの設定が最低条件を満たしているかどうかというのがとても大事になってくるので、PISA としてはシステム診断アプリケーションというのを作り、それを各国に配布し、予備調査と本調査の前に

3 番目の課題としては、生徒のコンピュータに対する習熟度のばらつき、これというのは、必ずしも私たちが測りたいと思っている、読解、数学、科学とは関係していないので、生徒のコンピュータに対する習熟のばらつきをできるだけ除いた生徒のコンピテンスを図りたいということで、練習のオリエンテーションを、テストの最初に提供しました。そして教科ご

課題(4)

- 経年変化の測定 - 筆記型調査(PBA)とコンピュータ使用型調査(CBA)では項目の難易度が変わる可能性がある。
→スケーリングにおいて、一部の継続問題を別問題として扱う。
- 学校/国によるコンピュータ性能の違い
→学校のコンピュータのハードウェアとソフトウェアの設定が最低要件を満たしているかどうかをチェックするシステム診断アプリケーションの提供
- 生徒のコンピュータに対する習熟度のばらつき
→練習のためにオリエンテーションを導入
- 多段階適応型テスト(MSAT)による非ランダムな項目割り当て
→等化された正答率(Equated P+)を計算する。
PISA 2022年調査から、適応型でない経路の比率を高くする。

単に計算することができなくなりました。

そこで、モデルを作って、それによって Equated P+ という、等化された正答率を特別に算出して、それによってこの問題は何パーセントの子供が正答したかを各国比較できるようにしました。

また 2022 年からは適応型ではないリニアな、ランダムに問題を当てはめる生徒の率も多くしまして、その生徒を元にして Equated P+ でなくても、普通に計算した正答率、P+ でも、各国比較できるようなデザインにしています。

課題(5)

- 経年変化の測定 - 筆記型調査(PBA)とコンピュータ使用型調査(CBA)では項目の難易度が変わる可能性がある。
→スケーリングにおいて、一部の継続問題を別問題として扱う。
- 学校/国によるコンピュータ性能の違い
→学校のコンピュータのハードウェアとソフトウェアの設定が最低要件を満たしているかどうかをチェックするシステム診断アプリケーションの提供
- 生徒のコンピュータに対する習熟度のばらつき
→練習のためにオリエンテーションを導入
- 多段階適応型テスト(MSAT)による非ランダムな項目割り当て
→等化された正答率(Equated P+)を計算する。
PISA 2022年調査から、適応型でない経路の比率を高くする。
- 公開問題の共有
→全言語の公開問題は、PISAウェブサイトにてご覧いただけます。

また、2018 年から適応型テストを導入したために、問題項目が生徒たちにランダムには割り当てられなくなりました。

それまで PISA というのは、全てのテスト問題項目が生徒にランダムに割り当てられていたのですが、2018 年から生徒の能力によって割り当てされることになるので、ランダムという前提が崩れました。それによって各項目の正答率というのを簡

公開問題の共有というのが、ペーパーベースより難しくなっていて、コンピュータの特別なプラットフォームがないと PISA のコンピュータの問題を示すことができません。幾つかの参加国がそうしたものを公表するキャパシティーがないということで、OECD 側で PISA の WEB サイトで全言語の公開問題を公開しています。

課題(6)

- 経年変化の測定 - 筆記型調査(PBA)とコンピュータ使用型調査(CBA)では項目の難易度が変わる可能性がある。
→スケーリングにおいて、一部の継続問題を別問題として扱う。
- 学校/国によるコンピュータ性能の違い
→学校のコンピュータのハードウェアとソフトウェアの設定が最低要件を満たしているかどうかをチェックするシステム診断アプリケーションの提供
- 生徒のコンピュータに対する習熟度のばらつき
→練習のためにオリエンテーションを導入
- 多段階適応型テスト(MSAT)による非ランダムな項目割り当て
→等化された正答率(Equated P+)を計算する。
PISA 2022年調査から、適応型でない経路の比率を高くする。
- 公開問題の共有
→全言語の公開問題は、PISAウェブサイトにてご覧いただけます。
- 今後進化する調査への適応

そして、最後には今後進化する調査への適用。PISA はどんどん進化していきまので、それに対応するのがなかなか難しい課題です。例えば、アクセシビリティ。障害のある子供たちをできるだけ PISA に入れていくというのが、一つの目標です。例えば目の見えない子供たちに、読んで聞かせる機能がコンピュータにはありますが、現在の PISA の問題は、左側

と右側の2つのパネルがあるので、それをコンピュータがどう読むかというのが課題になってくると思います。ですので、今後進化する調査へどう適用するかというのは大きな問題です。

ありがとうございました。

私たちの仕事について

詳しくは www.oecd.org/pisa をご覧ください。

テストを受ける: www.oecd.org/pisa/test

よくある質問: www.oecd.org/pisa/pisafaq

教育に関するPISA指標 GPS: <http://gpseducation.oecd.org>

PISA Data Explorer: www.oecd.org/pisa/data

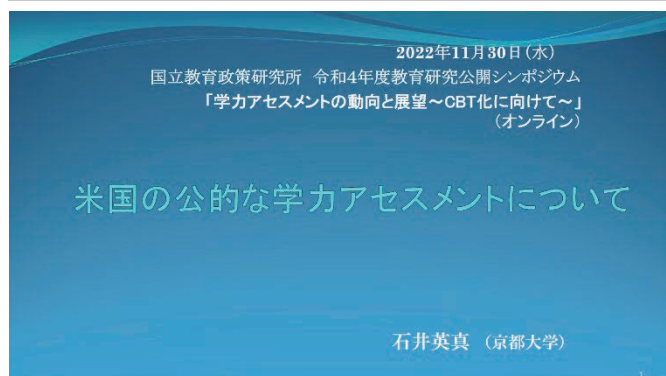
Email: Miyako.Ikeda@OECD.org



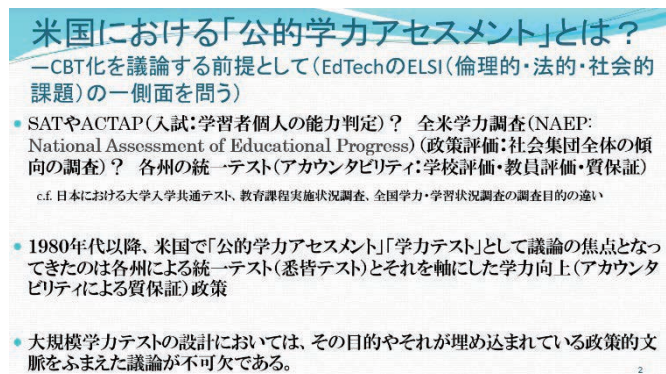
ありがとうございました。

②米国の公的な学力アセスメントについて

石井 英真：京都大学大学院教育学研究科准教授、国立教育政策研究所フェロー（事前録画）



石井：私からはですね、米国の公的な学力アセスメントについてというタイトルで報告させていただきます。



私の話は、直接 CBT 化の技術的な問題を明らかにしていくということよりも、むしろその前提となる、そもそもこの公的学力アセスメント、あるいは学力テストといったものは、一体どういうものなのかというあたりですね。その前提になるような議論を、アメリカの状況をお話しするという形になろうかと思います。ある種、CBT 化の ELSI についての検討ということになって

くるかなということですね。

まず、公的学力アセスメントであるとか、学力テストというふうに言った場合に何を指すのかということ。例えば、SAT とか、あるいは全米学力調査 NAEP、あるいは各州の統一テストなど、アメリカにおいては学力テストといったものもいろいろなものがあるわけですね。

しかし、今挙げたものは全て目的が違います。SAT は、これは入試ですから、学習者個人について能力を判定することになります。それに対し、NAEP というのは政策評価ですから、個人ということではなくて、社会集団全体の傾向を把握することになります。ですから、抽出調査が基本です。

全米の各州の統一テストというのは、これはアカウンタビリティが目的、質保証のためにやっていくということになりますので、個々の学習者よりも、学校単位の学校評価、あるいは教員評価が州単位で行われるということになります。

ですから、同じく学力テストと、あるいはアセスメントといった場合に、その目的によってその状況によって、大分要素は違ってくるということです。

例えば、日本でも全国大学入学共通テスト、教育課程実施状況調査、全国学力・学習状況調査それぞれ目的であるとか機能が違うと思いますね。

ですから、これらを一緒に議論してしまうと、いろいろ混乱の元になるということになります。よく、各国の比較というか、状況を調査した報告書等を見ますけれども、そのときに、国が違ったときにそれぞれの国の、全く違う目的のテストを比べているというものもたまに見たりします。改めて、学力テストというふうにいったときに、それが何を指すのかというあたりをち

やんと明確にしておかないと、状況を見誤ってしまうというふうなことになるはしないかなと思います。

大規模学力テストの設計においては、その目的であるとか、それが埋め込まれている政策的文脈、これを踏まえた議論が不可欠であろうということになります。ですから、限られた時間の中でお話するのは、アメリカにおいて、公的学力アセスメント、あるいは学力テストといったものが、どのように政策的に展開してきたのかというあたりですね、そこをお話するということになります。

特にアメリカにおいて、学力テストというのが注目を集めるのは1980年代以降ですね。特に各州における統一テスト、これ、悉皆ですけれども、それを軸にしたアカウンタビリティ政策の展開というのが、中心的な話題でありました。

全米的な学力向上政策の展開

1983年『危機に立つ国家』(連邦教育省長官の諮問委員会による報告書)の発表

＝アメリカの子どもの学力が危機的状況にあることを報告。
→州知事主導による改革が始まり、州の教育への関与が強まる。
学力向上が国家的な関心事となる。

1989年「教育サミット」の開催

＝ブッシュ大統領の呼びかけで全米の州知事が一堂に会して、「教育の国家目標(National Education Goals)」(ハイスクール卒業率の向上、数学・理科における世界最高水準の学力の達成などの6項目)を設定。

1991年『2000年のアメリカ:教育戦略』

＝2000年までに「教育の国家目標」を達成すべく、スタンダードの設定などの実行策が提案された。
→連邦政府のイニシアチブによる教育改革のはじまり。
州レベルでのスタンダード運動が全米規模で展開される。

それはこのスライドにまとめておりますけれども、80年代以降、要は州ごとに共通教育目標、アメリカは学習指導要領に当たるものがないので、それを全米で作ろうとしたけれども、できなくて州ごとに何とか作ろうよということになって、一応、州のスタンダードというのはできました。

しかし、アメリカは御存じのように分権的な国でありまして、それこそ州どころか

ですね、各学区であるとか、学校ごとにカリキュラムが大分違うのですかね。いかにスタンダードを作ったとしても各学区とか学校といったものがカリキュラム編成権を非常に強く持ちます。

ですから、参考資料ぐらいにしかならないわけですが、しかし、それに更に実効性を持たせるためにですね、州ごとにこのスタンダードにある種、準拠しながら、全てをフォローできているわけではないのですが、準拠しながら標準テストで、それで結果責任を求めていくというふうな、このスタンダードとアカウンタビリティによるシステムが整備されていくこととなります。

これは80年代以降、学力向上が全米的な課題になってくると言ったときに、実は、当時、日本は非常に経済的に状況がよくて、アメリカが日本に習えというので、日本には学習指導要領みたいなものがあるし、更にテストも結構やっているということで、それをアメリカ版でつくったのが、このスタンダードとアカウンタビリティのシステムで、えらく日本とは状況が全然違うわけですが、そういう形で展開していくこととなります。

2002年「一人の子どもも落ちこぼさない(No Child Left Behind:NCLB)法」の制定

1965年初等中等教育法の改正法であるNCLB法は、不利な条件に置かれている子どもとそうでない子どもとのギャップを12年間という期限付きで埋めることを目的として、補助金交付の条件として下記を各州に求めた。

- すべての子どもに適用されるスタンダードの設定、および、それに基づく州統一の学力テストの開発・実施(第3学年から第8学年までの各学年、第10学年から第12学年については、いずれかの学年で最低1回)。
- 各学区、学校に目標への到達度に関してアカウンタビリティを求め、その結果に応じて報酬や制裁を与える。複数年にわたって年間目標を達成できないと、その学校には段階的な措置が講じられ、最終的に教職員の人入れ替えもなされる。

2010年「共通スタンダード(COMMON CORE STATE STANDARDS)」の発表(全米州知事協会(National Governors' Association:NGA)と州教育長協議会(Council of Chief State School Officers:CCSSO)を中心に開発された)

言語技術と数学に関して、高校を卒業するとき生徒が学習しておくことが期待される内容や能力、および、小学校から高校の各学年において学習しておくことが期待される内容を示している。全米のほとんどの州で採択されている。

それが一番集約的に制度的に実装されて、更にいうといろいろと議論があったのは、この2002年のNCLB法というものです。

このNCLB法をきっかけに、ほとんどの州で、スタンダードとアカウンタビリティが整備されていくこととなります。

要は何かというと、各学校が、それぞれの状況に合わせて目標を設定します。学力

の達成度に関してね。それをちゃんと達成できていたらいいのですけれども、達成できない場合

ですね、最初は多少支援的なものもあるのですが、最終的にはですね、教員の総入れ替えみたいなかなり厳しいサンクションが与えられるということになります。

ですから、かなり結果責任を厳しく問おうという、そういう厳しい政策でありまして、これに対しては、様々に議論が起こって、このNCLB法というのは、非常に論争的な法律でありました。その後も、先ほど言いましたように、スタンダードとアカウントビリティを州ごとに整備していったとしても、各学校とか各学区それぞれ、非常にアメリカは多様な国ですから、カリキュラムは相変わらず多様なわけです。

それをさらに、州ごとのばらつきも少し是正していこうというので、共通スタンダードも作ったりとか、その中で、標準テストが暗記再生型の学力ばかり見ているから、もう少し思考力も含めたものを積極的に評価していこうという工夫もなされたところがあります。

けれども、逆に連邦政府が中心的にやっていくことに関しては、むしろトランプ政権などは、実は分権的にそれぞれが自由にやったらいいのではないかという考え方を持っていて、その辺で揺れ動いてきたというのがアメリカの状況であります。

ですから、このテスト政策が80年代以降のアメリカをかなり性格付けてきたというところがあります。

スタンダードに基づく教育改革 (standards-based reform) の世界的な展開

・教育の規制緩和や地方分権を進め、各地域や学校の創意工夫を促す一方で、学校教育に求める目標内容と達成水準(何をどの程度できるようになればよいか:スタンダード)を明確化し、学力テストなどによってその成果を検証したり、アカウントビリティ(説明責任)を果たさせたりしながら、学校教育の質を保証しようとする(遠隔統治構造)。

↑テスト準備教育をもたらすものであり、教育を単純作業化・規格化し、教育における競争主義と格差拡大を助長するものといった批判がなされてきた(佐貫・世取山, 2008)。

・学力テストの活用の仕方については、さまざまな形がある。アメリカのように、テスト結果による制裁が規定されていて、結果に基づいて教師や学校が厳しく値踏みされる国もあれば、オーストラリアのように、結果は公表しても支援的な介入を重視している国もあれば、イタリアのように、悉皆調査を行いつつ全国的な学力水準を確認するためにサンプルを抽出・分析し、あとは学校の自己評価に役立てさせる国もあれば、スウェーデンのように、学校間の評価の質の調整や学力観の共有に役立て、教師の自主的な活用を尊重している国もある(田中, 2016)。

そういった学力テストの結果によって結果責任を問うていく、ある種、成果主義的な、競争主義的な政策というのは、スタンダードに基づく教育改革などと言われて、これはアメリカだけではなくてですね、他の国でも多かれ少なかれ展開してきたところがあります。

例えばイギリスなどであれば、本当にその学校のテスト結果ですね、アメリカもそ

うですけど、そういったスコアを公開して、それで学校選択を促して、競争的な展開をしているところがあります。

これに関しては、結局、アメリカとかイギリスでそういうことをしていくのはなぜかという、学習指導要領がある日本においては、その中身をいじることによって教育現場に一定のインパクトを与えることはできるのですが、中身は実は連邦政府もいじれないんですね。各州も。そうしたときにそれぞれの学校現場とか、自治体の裁量を認めつつ、しかし、結果責任を問うことによって結果として誘導していくというある種、遠隔統治政策であったわけです。

しかしそうしてしまうと、結局のところ、テストしやすいもの、特にアカウントビリティの中心は標準テストでありましたので、そのテストしやすいものに学力観とか教育といったものを狭めてしまうのではないかと。

あるいは競争的で格差を拡大するのではないかとというふうに、様々な批判もあるわけです。

ただ、そのときに、学力テストの活用といっても、各国によって状況が大分違います。

アメリカのようにかなり政策がきついところもありますし、別の国はですね、割と支援的な介入が強いところもありますし、更に言うと悉皆ではあるけれども、基本的には、国は抽出して全体の傾向を把握して、あとは各学校でどうぞというふうにフリーハンドでらせていくということもあったりとかね。

あるいは、基本は学校ごとに評価をしていくのだ、しかし、ばらつきを少し是正するための参考資料ぐらいに共通テストを使っていくみたいな、様々な運用の仕方があります。

ですから、各国の比較をしていくときにも、質保証政策の文脈の中の学力テストの位置づけも、かなり各国によって状況が違うというところですね、この辺を踏まえた比較が重要になってくるのかなと思います。

標準テスト中心で競争主義的なアカウンタビリティ政策へのオルタナティブの模索

○標準テスト批判とパフォーマンス評価の誕生

=多肢選択などの客観テストは、現実世界と切り離された無味乾燥な文脈で、断片的な知識・技能を問うものである。交通法規や運転上の留意点をいくら知っていても、上手に運転できるかは実際に運転させてみないと分からないのと同じように、客観テストでは本物の学力は測れない。豊かに考える授業をしていながら評価では知識・技能(見えやすい学力)しか問われない、というミスマッチを解消し、目標・指導・評価の一貫性を確立する必要性。

●教室の外側のハイステイクスな標準テストによる、トップダウンで中央集権的で、学力競争を志向する、説明責任による目標・評価システムに対して、教室の内側の教師の判断を信頼する真正の評価と形成的評価による、ボトムアップで地方分権的で、学力・学習の保障を志向する、応答責任による目標・評価システムというオルタナティブも模索されてきた。

例) 米国における「真正の評価(authentic assessment)」をキーワードにした教育的なアカウンタビリティシステムの模索、「学習のための評価(assessment for learning)」をスローガンとする、米国のネブラスカ州、スコットランドやウェールズ、カナダのオンタリオ州などでの教室での教師の評価を尊重する評価システムの構築(石井, 2021aなどを参照)。

こういった標準テストを中心としたアカウンタビリティ政策に対しては、かなり批判もあって、それに対するオルタナティブも模索されてきました。

日本でも近年注目されているパフォーマンス評価というのは、もともと、そういったアカウンタビリティ政策に対する対抗的な動きの中で生まれてきたものです。要は、アカウンタビリティはかなり標準テ

ストによってなされることが多くて、その標準テストが測っているのは暗記再生型の学力で、現代社会が求めるような思考力であるとか、知識を使いこなしたりする力、そういったものは測れないでしょうと。

実際にそういった力は、その文脈の中でやらせてみないと評価できませんよねという話ですね。また、言うならば、車を運転すると言ったときに、実地せずにシミュレーションだけで、免許を与えないですよということと同じですから、やはり実際やらせてみないと、それは評価できないというふうな話から生まれてきました。

もう一つは標準テストみたいに外で評価の方法を決めるというのは、要は評価の方法というのは、結局目標であるとか、カリキュラムの中身を決めていくことになりまますので、事実上、現場のカリキュラム編成権を剥奪していくような動きにつながってくるわけですね。

それに対して現場でやっている教室の中で実際されていることを尊重するような評価、教室での評価をちゃんと尊重していきましょうよという、そういう動きがこのパフォーマンス評価の誕生には関係していたということがあります。

表. 学校改革の二つのアイデア (出典: Gallagher 2007, 29、石井, 2020.)

アカウンタビリティ	図存
企業モデル	民主主義モデル
学校改革	学校改善
改革の障害物としての教師	学校改善のリーダーとしての教師
一方通行の関係	相互関係
生徒の学業達成	生徒の学習
テストに基づく	豊かな情報に基づく評価
標準化	スタンダード
厳格な父親の道徳	共有された責任性の倫理
トランザクション	インタラクション
トップダウン	ボトムアップ(内から外へ)
純粋の行使	能力の構築
ハイステイクス	ハイインパクト
分不相応な不信	受けるに値する信頼
競争	協働
服従	関与
学習の評価(のみ)	学習のための評価(も)
単純さを要求する	複雑さを受け入れる

いずれにしてもですね、このスライドにもありますけれども、トップダウンのアカウンタビリティ政策に対して、ある種、草の根の、教室からの、参加型のといつかね、教室から立ち上げていくような改革の中でそういったきつ抗が、世界各国でいろいろ見られるということがあります。

こういった学力テスト政策の中で、様々な学力テストですとか、アセスメントが展

開してきているということがあります。

全米学力調査 (NAEP) (1969年開始)の変質

● 第一期(60~70年代) 政策立案に資するデータ供出ツールとして

Main NAEP (主調査:社会的要請の高い教育課題への取り組みを課題とし、最新の教育測定技術の適用と開発を行う)とLong-term Trend NAEP (動向調査:時系列的な学力変化を調べる)を定期的に実施。抽出調査で、生徒個人や学校単位のデータにはアクセスできず、州別でなく地域別に限定してデータを収集。

● 第二期(80~90年代) 全米学力調査自身に対する説明責任

連邦政府の教育予算に対する説明責任を果たす「より敏感な」ものにするべく、項目反応理論等の導入による調査デザインの洗練、および州別調査の解禁がなされた(スタンダードに基づくアカウンタビリティ政策との連結の始まり)。

● 第三期(2003年以降) 州の説明責任を果たすツールとして

NCLB法で州や学区が毎年作成する評価報告書において、州ごとに比較可能なデータとしてNAEPの州別調査のデータの使用が記されたことで、全州調査となった。

(木村, 2008の整理による)

調査する。

もう一つ動向調査というのは、経年比較でもって、全米の子供たちの状況がどうなのかということを見ていくということになっています。いずれにしても、両方とも政策評価としての調査的な役割が強いので抽出です。

更に生徒個人とか学校単位とか、州別とか、そういう個別の集計はほぼしないというように、そこはかなり禁欲的であったということがあります。ですから、そこがですね、目的によって、かなりデザインが違ってくるところです。

ただ、それが段々と、スタンダード運動の展開の中で性格が少し変質してきてですね、付け加わってきたというのが正しいかもしれませんが、2003年以降に関しては、各州がスタンダードとアカウンタビリティを進めていくときに、州ごとに立てているスタンダードを中身であるとか、難易度とかあるいは標準テストの難易度とかが、やはりばらつきがあるのではないかというふうなことに對して、このNAEPを使ったらその辺を調整できるのではないかという思わくがあって、州ごとに比較可能なデータを提供するというふうな役割を担わされるということにもなってきました。

そのように、基本的にアメリカのテスト政策というのは、各州のアカウンタビリティ政策、そこを軸にしながら、NAEPの在り方といったものも変わってきたということがあります。そういう文脈を押さえてアメリカのですね、テストであるとか、アセスメントは見ておく必要があるということですよ。

日本の全国学力・学習状況調査(全国学調)の性格

- 日本の全国学調は、改革や取り組みの成果(事実)を検証する役割や、出口(アウトプット)管理の方法としてよりも、実質的には、学習指導要領の学力観の趣旨を伝達し改革を先導・牽引する理念提示機能や、入口(インプット)管理を強化する機能を果たしている(石井, 2021b)。
- 全国学調実施前から、地方自治体レベルでも多くの悉皆の学力調査がなされてきた。しかし、一部の自治体を除けば、保護者や地域住民などへの、市町村別、学校別の学力調査の結果の公表は限定的である。結果の公表が認められる中、競争主義や成果主義を強化する方向で調査結果の利用が広まることには注意が必要である。ただ、英米と比べれば、現時点で日本の全国学調は必ずしもハイ・ステイクスなものではない。
- 日本において全国学調は、英米ほどには、教育の結果を厳格に管理・統制するものとはなっていない一方で、その結果によって上意下達の傾向や授業のスタンダード化が進行しないよう注意が必要である。

ますが、それはですね、アメリカとは少し状況が違うと思います。

日本の全国学力調査というのは基本的には、アウトプット、アウトカムのところでは結果責任を問うというよりも、むしろ学習指導要領の学力観であるとか、趣旨といったものを、テスト問題を通じて伝達していく。B問題、活用問題なんていうのが一番端的にあらわれていますよね。

そういうふうに、理念提示機能が非常に強かったと思います。ですから、日本において学力調査というのは、英米ほどには教育の結果を厳格に管理統制するものとはなっていない。しかし、その一方で、むしろそれ以上というかですね、教育の目標とか内容及び実践の方法を特定の在

基本的にアメリカにおいては、その州の統一テストが中心的なトピックになってきたわけですが、このNAEPというのはそれほど注目される調査ではないです。

ももとは政策立案のための政策評価として作られたものですので、伝統的に主調査というのは、その時代時代の教育課題に対して、実際、どうなのかということ

一方で、日本はどうだったのか、どうなのかということですが、日本の全国学力調査と、アメリカの競争的なある種、結果責任を問うようなテスト政策といったものを同列に扱って、日本もアメリカみたいに競争主義とか結果責任を問うような厳しい状況になっているのではないかという論調がありますけれども、確かに管理統制的な側面は強まっているとは思

り方へと、より直接的に枠づけるという形で機能しているかなと思います。各自治体レベルは結果言われるわけですね、ランキングが出されますからね、その結果、現場自体はランキングが出されないということがあって、それで自由になっているかといったら、実はそうではなくて、その自治体ごとに学力向上策なんてものがパッケージとして組み立てられて、それが現場において推進されていきます。その結果として、上意下達の傾向であるとか、「このとおりのやれよ」みたいな形で、授業改善のある種、指針みたいなものですが、それがマニュアルみたいになるものですから、現場の創意工夫がすごく萎えてしまって、むしろその箸の上げ下ろしみたいなものがすごく規制されるというふうな状況になっている可能性はあるのかなと思います。

ですから、学力テストの問題を CBT 化していくというときにも、そうすることによって、今の全国学力・学習状況調査の状況がどういうふうに変わっていくのかということも踏まえながら議論していくことが必要になってくるのかなと思います。

CBT化に向けた論点提示

- テストの目的によって CBT 化や項目反応理論の導入の可能性や望ましさは変わる。特に、項目反応理論の導入に際して、調査問題が原則的に非公開になる点、問題バンクの作成やデータ管理のコスト、および、問題漏えいのリスクなどが大きな課題となる。過去問で勉強するなどの日本のテスト文化との関係、あるいは、学習指導要領の改訂のたびに問題バンクを変えることになるのかといった日本の教育制度との整合性も検討が必要だろう。
- 調査的性格が強くてハイ・ステイクスな性格が弱い、経年変化を見る補完調査、あるいは児童生徒や教員や学校への質問紙調査等への利用はしやすいだろう。逆に、入試への導入はハードルが高い(木村, 2022)。また、悉皆の本体調査も、問題公開の難しさや CBT 化しやすいものに学力観が限定されることなどをふまえると、理念提示機能、および個々の学校・教師・児童生徒の指導・学習改善機能において工夫が必要だろう。全国学力・学習状況調査の主目的を絞り、個々の学習者の支援のあり方等についても検討したりしながら(MEXCBT 等で別途フォロー)、さらに、学力の質と評価方法との対応関係も考えながら、技術的制約によって目的や中身が規定されないよう注意が必要である。

改めて、CBT 化に向けた論点提示ということになりますけれども、CBT 化あるいは CBT の旨味のひとつとしては、複数回実施であるとか、実施時期に柔軟性を持たせられるということになるわけですが、その前提としては項目反応理論の導入があるだろうということです。

しかし、技術的な導入の可能性とか望ましさといったものは、目的によって大分変

わってきます。

例えば、項目反応理論の導入に関しては、様々に議論されているところですが、調査問題が原則的に非公開、あるいは問題バンクの作成とかデータ管理の方、こちらの方にコストが取られるから、確かに問題の印刷とかそういったもののコストはなくなりますけれども、その問題バンクのですね、作成とかデータの管理とか、この辺のコストというのは結構大きいわけですね。

リスクとしては問題漏えいのリスク、これがかなり大きいと。

例えば入試であれば、これまでだとカンニングというのが不正行為なのですが、問題漏えいですね。1問でも外に漏れてしまうということになってくると、これ大問題になるわけですね。ですから、この辺をかなり厳密に管理していくというのは相当な工夫であるとかコストがかかってくるというあたりは、ちゃんと考えておかなければいけないということですね。CBT 化と言ったときに単にコンピュータで入力するというふうな話ではなくて、システムを大きく変化させていくということが背景にある。

ここと現行の状況との整合性をどう考えていくのかということになってくるかなと思います。

例えば、過去問で勉強するといった日本のテスト文化との関係はどうなのか、あるいは問題バンクをまとめて作ると言ったときに、学習指導要領の改訂のたびにですね、問題バンクを変えるみたいなことにつながるの、この辺をどう考えていったらいいのかというあたりですね。現行の日本の教育文化であるとか、あるいは教育制度との間の整合性、こういったところも考えていく必要があるのかなと思います。

そういったときに、調査的な性格が強くて、ハイ・ステイクスな性格が弱い場合はある種、CBT 化もしやすいのかなということですね。CBT 化に向けたワーキンググループの報告書、この最終まとめを見たりしますと、主調査とそれから補完調査ということで、経年変化の方が補完調査なのですが、こちらは調査的な NAEP に近いような形で政策評価的側面が強いので、ですから、

児童生徒とか教員や学校への質問紙調査等ですね、こういった調査的なものに対しての利用は考えやすいかなと思います。

逆にですね、やはり入試への導入というのは、かなりハードルが高いのではないかなというふうに思います。

また、全国学力・学習状況調査のこの本体調査というふうに言われているものですが、この悉皆の調査に関しては、様々に検討する必要があるのかなと思いますね。

問題公開の難しさというもの、あるいは CBT 化しやすいものに学力観が限定されるのではないかなというのは、もともと、悉皆調査でやってきた全国学調の理念提示機能、あるいは全員にやることによって個々の学校とか教師とか児童生徒のその指導とか学習の改善機能、この辺との関係がどうなのかというあたりは、ちゃんと検討が必要なのかなと思います。

やはり、この理念提示機能ということで、調整が大分必要になってくるかなと思いますし、個々の学習者の支援ということに関して言うと、やはり調査問題が非公開になってくると、もう何か模試みたいなデータだけでですね、それで改善のための十分なデータとなかなか言えないかもしれませんね。やはり調査問題があって、個々の子供たちと照らし合わせることによって、つまりきがよくわかってくるということもあると思いますので、そうすると直接的に全国学力調査の主目的の中にそれを入れるべきなのかどうか、この辺、主目的をちゃんと絞ってですね、それで悉皆調査であることの是非といったものも、やはり再検討をしながら何ならこの補完調査みたいなものの目的こそが主目的になるかもしれませんけれども、その辺ですね、調査としては目的も検討しながら進めていくということが重要かなと思います。

方法が中身を規定していくということは多かれ少なかれ起こるわけですが、その辺ですね、CBT 化ということによって、どういう学力が見えて、どういうことが逆に見えなくなるのかと。CBT 化によって確かにシミュレーションみたいな問題文をだらだらだらだら書くのではなくて、状況的なシミュレーション的なものというのは、割とやりやすくなるかもしれないですね。

しかし、それも先ほど申しましたように、実地で実際に車を運転するというをせずに、単にシミュレーションだけで運転したことにするかどうかというのは、これはまた話が違いますね。だから改めてその目指す学力の質といったものと評価方法との対応関係、この辺をまた検討していくことが必要かなと思います。

繰り返しになりますが、この CBT 化といったものは、単にコンピュータで入力するという、単にデジタル化するというだけではなくて、技術的に様々なシステムの変更を伴うと。ですから、そのシステムの変更による、この技術的な制約によって、そもそもの目的であるとか、中身といったものが過度に規定されてですね、それによってそもそもの目的が果たせないというふうなことにならないように、そこは冷静な議論が必要なのかなというふうに思います。

では以上で終わらせていただきます。

引用・参考文献

- 荒井克弘・倉元直樹編著(2008)『全国学力調査 日米比較研究』金子書房。
- 石井英真(2015)『教育実践の論理から『エビデンスに基づく教育』を問い直す—教育の市場化・標準化の中で』『教育学研究』Vol. 82, No. 2.
- 石井英真(2017)『学校改革とカリキュラム変革の系譜』『岩波講座 教育 第五巻 学びとカリキュラム』岩波書店。
- 石井英真(2020)『再増補版・現代アメリカにおける学力形成論の展開—スタンダードに基づくカリキュラムの設計』東信堂。
- 石井英真(2021a)『カリキュラムと評価の改革の世界的標準化と対抗軸の構築』広瀬裕子編『カリキュラム・学校・統治の理論』世織書房。
- 石井英真(2021b)『学力調査を問い直すための評価リテラシー』西岡加名恵・石井英真編著『学力テスト改革を読み解く—「確かな学力」を保障するパフォーマンス評価』明治図書。
- 川口俊明(2020)『全国学力テストはなぜ失敗したのか—学力調査を科学する』岩波書店。
- 本村拓也(2008)『2003年以降の全米学力調査の変遷』荒井克弘・倉元直樹編著『全国学力調査 日米比較研究』金子書房。
- 本村拓也(2022)『新しい情報技術と大学入試—大学入試のCBT化を巡る『試論』が導き出すパラドクス』石井英真・仁平典宏・濱中淳子・青木栄→丸山英樹・下司編著『教育学年集13 情報技術・AIと教育』世織書房。
- 佐藤仁・北野秋男編著(2021)『世界のテスト・ガバナンス』東信堂。
- 佐賀浩、世取山洋介編(2008)『新自由主義教育改革—その理論・実態と対抗軸』大月書店。
- 田中耕治編(2016)『グローバル化時代の教育評価改革—日本・アジア・欧米を結ぶ』日本標準。
- C. W. Gallagher, 2007 *Reclaiming Assessment: A Better Alternative to the Accountability Agenda*. Portsmouth, NH: Heinemann.

第二部

パネルディスカッション これからの学力アセスメントーCBT化に向けて

コーディネーター

白水 始：国立教育政策研究所 初等中等教育研究部副部長・総括研究官

パネリスト

巖岩 晶：国立教育政策研究所 教育データサイエンスセンター総括研究官

渡辺 洋平：埼玉県教育局市町村支援部義務教育指導課長

佐藤 寿仁：国立大学法人岩手大学教育学部准教授、
国立教育政策研究所 CBT スーパーバイザー

池田 京：OECD 教育スキル局シニアアナリスト

国立教育政策研究所 令和4年度教育研究公開シンポジウム
「学力アセスメントの動向と展望～CBT化に向けて」

パネルディスカッション
「これからの学力アセスメント
～CBT化に向けて」

■OECD 教育スキル局 池田 京シニアアナリスト
■埼玉県教育局 市町村支援部 渡辺洋平義務教育指導課長
■岩手大学 教育学部 佐藤寿仁准教授
■国立教育政策研究所 教育データサイエンスセンター 巖岩晶総括研究官
■国立教育政策研究所 初等中等教育研究部 白水始総括研究官



白水：皆さん、こんにちは。御紹介にあずかりました白水でございます。パネリストの先生方、どうぞよろしくお願いします。私からは、まずごく簡単にパネルの趣旨と構成をお話しします。このパネルでは、先ほどの石井先生の論点提示にあったように「CBT化ありき」ではない議論、これからの学力アセスメントはどうあるべきかという議論を行いたいと考えます。

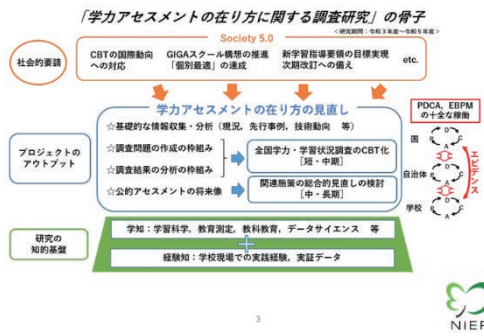
ここまでの振り返り

- ・基調講演（大津先生）
CBT化に際しての二本柱による役割分担の構想
・教育政策評価を主目的とした経年調査
・指導改善を主目的とした悉皆調査
- ・講演（池田先生・石井先生）
・国際調査におけるCBTの導入、その利点や可能性
・調査目的に合わせた調査設計の必要性

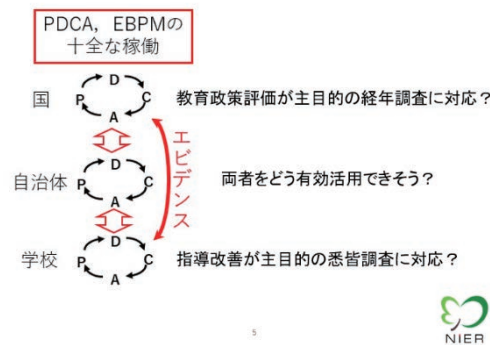


まず、ここまでの振り返りです。大津先生の基調講演では、CBT化に際しての2本柱による役割分担、教育政策評価を主目的とした経年調査と指導改善を主目的とした悉皆調査、この役割分担が大事ではないかという話でした。石井先生、池田先生の講演からは、調査目的に合わせて調査設計をすべきこと、その目的に合わせたときに、CBT導入によって、どのような利点や可能性が広がるかということをお紹介いただきました。

「なぜ目的に合わせた設計が必要なのか」をもう一步踏み込んで考えてみると、それは私たちの取組や実践ーそれが教育政策の評価であれ、学習指導であれーその取組が果たしてうまくいつているかどうか、これまでの取組を振り返って、明日からの取組を良くしていくため、そのために評価というのがあるからだ。そんなふうと考えられるのではないかと思います。

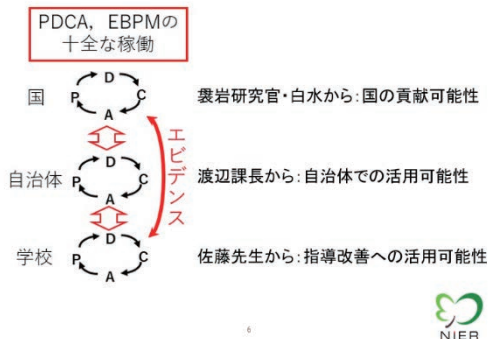


そこで、国研の学力アセスメントの調査研究では、最初の趣旨説明にあったように、右側のサイクルで示しておりますが、国・自治体・学校、それぞれのレベルでそれぞれの取組・実践のPDCAサイクルを充実させていくための学力アセスメントは、どうあるべきか、ということを検討しております。

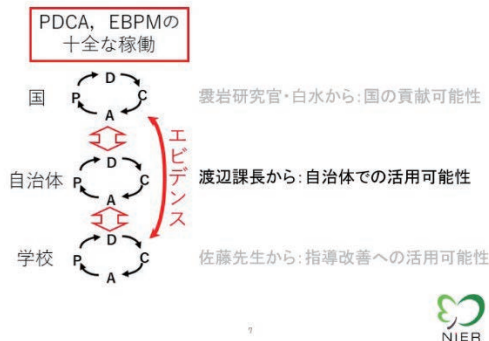


この部分をもう少し大きくしました。国のPDCAのところ、先ほどの2本柱に対応させますと教育政策評価ということになりますので、これが経年調査に対応して、そして指導改善が主目的の悉皆調査が学校のPDCAサイクルというのに対応しように思います。けれども、この両者というのは完全に分離したものなのか。それとも結び付けて使っていいのか。そのとき、自治体というのは一体どういう役割を果たしていけるのか。そんな検討ができると、今回学校現場の先生

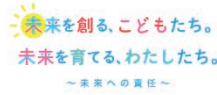
方、そして教育委員会の指導主事の先生方がたくさん参加されておられますので、アセスメントを実践的に考えていくヒントになってくるのではないかと。



そのために、このパネルでは、まず渡辺課長から自治体で全国学調というのをどういうふうに関用できるか、あるいはそれをもう県で独自でなさっている取組を紹介していただいた後で、佐藤先生から学校の指導改善への活用可能性、そして、斐岩研究官と私の方から国の貢献可能性というお話をしながら、池田先生から国際的な観点で御助言を頂くような展開をしたいと考えております。



それでは、早速ですけれども、渡辺課長から埼玉県という一つの自治体が行っている調査をお話しいただいて、そこにCBT、あるいは全国学調の役割含めてお話しいただければと思います。それでは、渡辺課長、報告をよろしくお願ひします。



国立教育政策研究所 令和4年度教育研究公開シンポジウム

「埼玉県学力・学習状況調査」について



令和4年11月30日（水）
埼玉県教育局市町村支援部
義務教育指導課長 渡辺 洋平

紹介させていただければと考えてございます。

1 埼玉県学力・学習状況調査の概要（平成27年度から実施）	
実施目的	児童生徒の学力や学習に関する事項等を把握することで、教育施策や指導の工夫改善を図り、児童生徒一人一人の学力を確実に伸ばす教育を推進する
調査対象	調査を希望する県立及び県内市町村立の小・中学校等に在籍する小学校第4学年から中学校第3学年の全児童生徒（R4年度：62市町村・約27万人）
調査事項	①児童生徒に対する調査 【教科に関する調査】 小学校第4学年から第6学年まで 国語、算数 中学校第1学年から第3学年 国語、数学、英語（第2学年から） 【質問紙調査】 学習意欲、学習方法及び生活習慣等に関する事項 ②学校及び市町村教育委員会に対する調査 学校における教育活動並びに学校及び市町村における教育条件の整備等に関する事項
特長	学力の伸び（経年変化）などを継続して把握することのできる調査 ・小学校第4学年から中学校第3学年までの同一児童生徒を継続して把握 ・PISA（国際学力到達度調査）と同様の調査手法（項目反応理論）を採用
県外における実施状況	108市町村で調査を実施 → 約14万人の児童生徒が受検 ＜実施自治体＞福島県（全域）、高知県、鳥取県、滋賀県、福岡県（一部） 広島県福山市、宮城県白石市、高知県梶原町、東京都三鷹市、鳥根県益田市（全域） 神奈川県秦野市、小田原市、大阪府堺市（一部）

渡辺：はい。今、御紹介にあずかりました埼玉県教育局義務教育指導課の渡辺でございます。本日はよろしくお願いたします。資料を共有させていただきます。

先ほど御紹介いただきましたとおり、私の方からは現在、埼玉県で取り組んでおります埼玉県学力・学習状況調査につきまして、その概要、また、その活用、そして現在進めております CBT 化に向けた取組を

まず、概要についてです。埼玉県学力・学習状況調査につきましては、平成27年度から実施しております。令和4年度で8回目でございます。

目的については記載されているとおりでございますけれども、児童生徒の学力や学習に関する事項を把握し、教育施策や指導の工夫改善を図り、児童生徒一人

一人の学力を確実に伸ばしていくといったことを目的として実施してございます。

調査対象につきましては、県内の政令市であるさいたま市を除く62市町村全体、小学校4年生から中学校3年生までの合計約27万人が直近の令和4年度は調査を受検している状況でございます。調査事項につきましては、教科について国語、算数・数学、中学校2年生3年生については英語を実施している状況でございます。

また、教科に関する調査だけではなく、質問紙調査で学習意欲でございましたり、児童生徒の学習方法、生活習慣などについても、例えば直近であればコロナと学習態度についても、クロスしながら調査している状況でございます。

本調査の最大の特徴につきましては、今までの先生方の講演でもございましたとおり、学力の伸び、経年変化を継続して把握することのできる調査ということでございまして、小学校4年生から中学校3年生まで同一の児童生徒を継続して、その伸びを見ていくといったことで、IRT、項目反応理論を採用した調査で実施してございます。

また、連携・協定等を結びまして、埼玉県の約27万人以外の児童生徒以外にも福島県でございましたり、高知県の一部など、108市町村14万人が協定を結びながら、この調査を実施している状況でございます。

2 埼玉県学力・学習状況調査の特長（学力の伸びが明確に示される）

※ 本調査では、学力を「学力のレベル」として提示している。
一小学校第4学年から中学校第3学年までの全ての問題に難易度を設定することで、学力のレベルを測定している。

○ 本調査における「学力のレベル」の考え方
・「どのくらい難しい問題を解く力があるか。」を学力のレベルで表している。レベルが上がると、難易度の高い問題を解く力がある。

○ 本調査における「学力の伸び」の考え方
・年度間の「学力のレベルの差」を学力の伸びと捉える。

○ 学力のレベルは、36段階（12レベル×3層）で設定しています。

2

続きまして、特長でございますが、学力の伸びが明確に示されることで、小学校4年生から中学校3年生までの全ての問題に難易度を設定してございます。項目反応理論の基本でございますけれども、難易度を設定することで、学力のレベルを測定することができ、右側の表のように、学力のレベルを36段階で設定した上で、経年で一人一人の伸びを丁寧に見ていく特長でございます。

2 埼玉県学力・学習状況調査の特長（学力を伸ばした取組や指導方法の共有）

・学力を良く伸ばしている学校が分かる

・学力を伸ばした取組や指導方法を共有し、一人一人の学力の伸びにつなげる

・学力を良く伸ばしている学級や教科が分かる

前年度のクラス	「学力が伸びた児童」の割合	国語	算数
5-1		73.4%	96.5%
5-2		91.8%	75.1%

○ データに基づく学校・教員の授業改善PDCAサイクルの確立

3

また、特長として、継続的に見ていくということで、学力を伸ばした取組でございましたり、指導方法の共有にも力を入れてございます。左上の方にございますけれども、この調査によって、学力をよく伸ばしている学校が見てとれます。

それはすなわち、学力が高い学校では必ずしもなく、学力が高くて伸び悩んでいる

学校があったり、あるいは学力が低くてもよく伸ばしているという学校があり、そういった各学校・学級の実態がよく把握できる特長をもってございます。

また、左下の方にございますけれども、学力をよく伸ばしている学級あるいは教科がこの調査を通じて見てとれます。

資料に例をあげましたけれども、前年度の5年2組は、国語がすごくこの伸びているといったような状況が、この調査から毎年毎年見て取れる状況でございます。こういったデータから各学校において、又は各学級において、学力を伸ばした取組、あるいは指導方法といったものの共有が可能となり、一人一人の学力の伸びにつなげております。右側の方にございますけれども、前年度に学力を伸ばした学校や学級の取組、教科の指導方法を把握した上で、調査結果を分析して、新しい年の取組、指導に生かしてもらおうという日々の指導改善を行うことで、児童一人一人の学力向上につなげていくといったことございまして、いわゆる肌感覚といったものではなくて、データに基づきながら、学校・教員の授業改善のPDCAサイクルを確立している特長がござい

3 埼玉県学力・学習状況調査を活用した学力向上施策

(1) 県の取組

埼玉県

総合的な支援

- 各市町村・学校における取組共有の促進
- 調査データの分析結果の普及
- 効果的な取組を県内で広く共有

重点的な支援

- 特に支援が必要な市町村・学校への対応
- 特に学習のつまずきが見られる学年・教科等への対応

市町村

学校

児童生徒一人一人の学力向上

Plan 計画を立てる

Do 取り組みを行う

Check 学力調査結果を活用する

Action 取組を改善し、次につなげる

施策名

- 良い授業を見つけ！広め！学力UP事業
- AIを活用した学びの実践研究事業
- 学力向上推進協議会
- 県学級活用研修事業

施策名

- 「未来を生き抜く人財育成」学力保障スクラム事業
- 県学級の結果に基づく市町村や学校に対する支援
- 学力向上研究校指定事業
- 学力向上プロジェクト教員の配置

4

続きまして、埼玉県学力・学習状況調査を基礎とした県の取組、学力向上施策についても、簡単に御紹介させていただければと思います。県といたしましては、2本柱で学力向上の支援を実施しているところでございます。まず、左の方にございます、総合的な支援ということで、各市町村学校における取組、よい取組の共有の促進ということで、いわゆる好事例の横展開を実施

してございます。また、2つ目の柱として、重点的な支援ということで、埼玉県学力・学習状況調査、あるいは全国学力・学習状況調査などから、データから見て取れる特に支援が必要な市町村、あるいは学校への対応、あるいは特に学習のつまずきが見られるような学年・教科等への対応に重点的な支援という形で、ピンポイントで支援を行っていくといった取組を行ってございます。

3 埼玉県学力・学習状況調査を活用した学力向上施策

(2) 主な施策

施策名	事業コンセプト	成果
総合的な支援	良い授業を見つけ！ 広めて！学力UP事業	県学調の結果から学力を伸ばしている教員の特別付きの取組資料を作成。学力向上に効果的な授業等のノウハウを共有・普及する。
	AIを活用した 学びの実践研究事業	県学調の結果をAIで活用し、児童生徒一人一人に合わせた指導及び最適な学びを実現する。
重点的な支援	「未来を生き抜く人財育成」 学力保障スクラム事業	家庭の経済状況などが学力に課題を抱える小学校4・5年生の児童に対する教育的支援の方法を実施校で研究する。 【実施校：10校（令和3年度）】
	超学調の統廃しに基づく 市町村や学校に対する支援	前年度の県学調で「学力が伸びた児童生徒の割合」が県平均を下回るなどの状況にある市町村や小学校の調査を実施する。 【実施市町村：8市町・実施校：16校（令和3年度）】

5

上で、解説をつけて、映像資料を作成した上で、県の研修あるいは市町村の研修で活用してもらう形で学力向上に効果的な授業のノウハウの共有・普及を図ってございます。また、重点的な支援の例として、一つあげさせていただくと、上から、3つ目の「未来を生き抜く人財育成」学力保障スクラム事業」を実施しております。就学支援の状況でございましたり、県の学力調査の状況を踏まえまして、学力に課題を抱える小学校に対して、特に小学校4年生5年生の児童に対して、支援を行う取組を県として実施しております。このように県では総合的な支援、重点的な支援の2本柱で学力向上に係る取組を進めているところでございます。

4 埼玉県学力・学習状況調査のCBT化

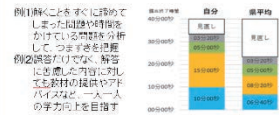
(1) 目的

※CBT: Computer Based Testingの略で、コンピュータを使用して実施する調査
これまで紙の調査で取得してきたデータに加え、より精緻なデータを取得、蓄積することが可能となり、エビデンスに基づいた施策の検討、教員の指導方法の工夫・改善を進め、児童生徒一人一人の更なる学力向上を目指す。

(2) 主なメリット

①より精緻なデータの取得、蓄積

正誤の状況に加え、解答時間分析することで、児童生徒のつまずきを把握し、より適切な指導が可能



②結果返却までの時間短縮

自動採点により結果返却までの期間短縮により、学校における児童生徒への指導改善が促進・向上

③CBTの特性を活かした出題

児童生徒の学習をより引き出すような出題や、映像を活用してより実際の学習場面を再現した出題



④学校の負担軽減

学校における調査資料(問数標準等)の受け取り、回収のための箱入れ等、調査資料の管理が不要

⑤コスト削減

調査資料の印刷、配達及び回収等の費用が不要

6

ございまして、1人1台端末の活用が可能になったといったことも踏まえて、CBT化に今、舵を切っているところでございます。主なメリットといたしましては、先ほど、先生方からも御紹介がありましたけれども、1つ目といたしましては、より精緻なデータの取得が可能になるといったことで、今までは〇×（正誤）の状況しか見てとれなかったものが、CBT化することによりまして、解いた時間であったり、そのようなログが、新たに取れるようになりますので、そういったものを指導に生かす。具体的には、例(1)のように、解くことをすぐに諦めてしまった問題、あるいは時間をかけている問題を〇×と併せてクロス分析させながら、つまずきを把握して指導に生かすことを進めたいと考えてございます。また、②でありますように、結果返却までの時間短縮に向けて自動採点などの取組を行うことによって、できるだけ早く学校に児童生徒にフィードバックをしたいと考えているところでございます。また③にありますようにCBT化の特性を生かした出題が可能となります。プレ調査でも実施しましたが、実際の学習場面に即した出題が可能になります。2つの例にある通り、例えば吹き出しを使用して会話を再現した上で、それについ

具体的な例を御紹介させていただければと思いますが、時間の関係上、それぞれ一つずつ御紹介させていただきます。まずは総合的な支援という横展開として、一番上でございます「良い授業を見つけ！広めて！学力UP事業」を実施しております。県学調の結果から、学力を数年にわたって伸ばしている先生が見てとれます。そういった先生のいわゆる良い授業を撮影した

最後に、今日の本題でございますCBT化に関して、御紹介させていただければと思います。埼玉県学力・学習状況調査は、8年目で一人一人の学力の伸びを着実に見ていくことで、一つの指標ではございませけれども、全国学力・学習状況調査の状況などからも学力向上の成果が現れていると考えてございます。他方でこの取組をより一層加速させたいと県としては考えて

ての設問を設けたり、あるいは算数・数学であれば、動画の出題が可能となります。コンパスの作図を見た上で、それについての問題を解くといったような実際の学習場面に即した出題が可能になると考えてございます。

4 埼玉県学力・学習状況調査のCBT化

(3) 取組の概要

様々な課題に対応しながら、段階的に埼玉県学力・学習状況調査(県学調)のCBT化に向けた取組を推進

R3年度	① 試行調査 【目的】3種のOS(Windows・Chrome・iPad)で、直接型・筆約型のどちらの接続方法でも調査実施が可能などの確認。 児童生徒のICT活用能力の確認 【対象】4市町(各小・中学校1校、私立中学校1校)
R4年度	② 予備調査 【目的】「学力の伸び」をCBT化後も継続して測定する仕組みの構築 【対象】4市町(各小・中学校1校)
	③ 接続確認調査 【目的】私立中学校及び県内(さいたま市除く)の全小・中学校を対象とした、学校の通信環境(一斉接続)の確認 【対象】62市町村全校(約1,000校)及び私立中学校(1校)
	④ プレ調査 【目的】全県実業を見据えた実施方法や結果返却方法の実証 【対象】4市全校(小学校7校・中学校3校)
R5年度	⑤ 県学調(本調査) ・県内の市町村(さいたま市除く)がCBT化による調査(従来通り)を実施し実証予定 ※CBT(36市町村、約12万人)、紙による調査(26市町村、約17万人)
R6年度	⑥ 県学調(本調査) ・私立学校及び県内全ての市町村においてCBTで実施を目指す

7

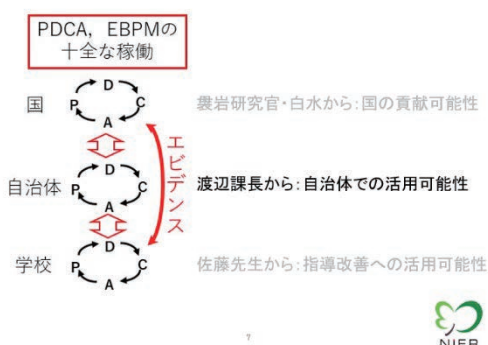
他方で、CBT化に向けては、様々な課題があると考えてございまして、その課題一つずつをクリアしながら、現在、その取組を進めているところでございまして、令和3年度から一歩ずつ進めておりまして、この表を見ていただくとおわかりのとおり、例えば、令和3年度、試行調査では、県内3種類のOS(Windows・Chrome・iPad)がござ

いますので、そういったものできちんと実施できるのかどうか、また、学校の通信環境についても、通信環境が学校からの直接型もあれば、市町村のサーバを通るもの、いろいろ形がございまして、そういった接続方法で可能かどうか、そういったことをまず丁寧に確認する作業を実施しております。

今年度につきましては、予備調査で学力の伸び、この伸びを見ていくことが、埼玉県学力・学習状況調査の最大の特長でございまして、それが継続してできるのかどうかといった確認作業を実施しました。一番大きな課題の一つとして、学校の通信環境がございまして、学校で3学年全体で使ってしまうと、どうしてもうまくネットに接続できないといった状況も見て取れましたので、そういった通信環境の調査を一部の市町村ではなく、埼玉県内全小中学校を対象として、実施したところでございまして。

こういった取組を一步一步、現在進めているところでございまして、ようやく来年度につきましては初めて本調査を一部の市町村、約半数の市町村においてCBTで本調査を実施いたします。その上で、令和5年度の結果を見つつではありますけれども、令和6年度には、埼玉県学力・学習状況調査の本調査を全面的にCBT化で実施したいと考えているところでございまして。埼玉県といたしましては、児童生徒一人一人の学力向上に向けて、市町村や学校の声も丁寧に聞きながら、試行錯誤をしながら取組を進めているところでございまして、国のお力も借りながら、引き続き進めていきたいと考えているところでございまして。

私からのプレゼンは以上でございまして、ありがとうございました。



白水: ありがとうございました。学力の伸びを継続的に把握したいということで、全国にも先駆けてIRTを導入されて、それをPBTでも実現されていたのですが、それを更に加速しようとCBTを始めておられるというお話だったかと思います。

それでは、池田先生、今のお話聞いていただいた上で、教育施策の評価、あるいは指導改善の点で、御感想やコメントをいただければと思います。

池田：ありがとうございました。埼玉県は本当にリッチな、ものすごく有効なデータを持っていらっしゃるというのがとてもよくわかりました。一つ質問なのですが、この教育のデータを、他の公的なデータと結び付けることができるのでしょうか。プライバシーの問題もありますし、難しいことではあるのですが、より生徒を個人として全体を見るためにも、そうしたほかの教育以外のデータとつなげることが、もしかしたら有効かもしれないと思いました。

また、PISA では生徒の成績だけではなくて、公平性（エクイティー）というのをよく見ているのですね。生徒のバックグラウンド、家庭背景と成績がどれだけ関係しているか、その辺も見ていらっしゃるのかなど。生徒の成績の伸びだけではなくて、エコノミーバックグラウンドとのギャップがどれだけ経年で増えているのか減っているのか。理想としては、どのような背景を持っている子供たちもいい成績がとれるというのが理想ですので、それに近づいていらっしゃるか、経年的に見られるのかというのが一つ質問です。

最後の質問ですが、PISA で、今、一つの議論になっているのが、教育の支援ってというのが、必ずしも全ての生徒、学校に均一に効果的になるわけではなく、文脈の違いによってあるタイプの支援はとても有効であったり、ある学校に対しては有効であるけれども、他のタイプの学校に対しては有効ではない。もしかしたら、これは生徒のレベルでそれもあり得るかもしれません。埼玉県はすごくデータを持っていらっしゃるので、統計を使って、教育政策の効果の不均一性というのを見ていけるのではないかと思います。

白水：ありがとうございました。非常に重要な質問でしたので、渡辺課長、答えられるものだけで結構でございます。どうぞお願いいたします。

渡辺：ありがとうございます。様々なデータとつなげていくといったことについて、我々としても重要と考えています。現に、県内の市町村によっては、いろいろなデータと埼玉県学力・学習状況調査のデータをひも付けながら生かしていることもございます。他方、県全体としては、やはりそのデータの取扱いの問題であったり、まだ課題があり、更に検討を進めていかないとはいけません。引き続き、メリット・デメリットを整理しながら、そこは進めていきたいと考えてございます。

2点目についても非常に有り難い意見だと思いますので、今後、前向きに考えさせていただければと思います。

PDCA, EBPMの
十全な稼働

国
P → D → C → A → P
豊岩研究官・白水から：国の貢献可能性

自治体
P → D → C → A → P
渡辺課長から：自治体での活用可能性

学校
P → D → C → A → P
佐藤先生から：指導改善への活用可能性

エビデンス

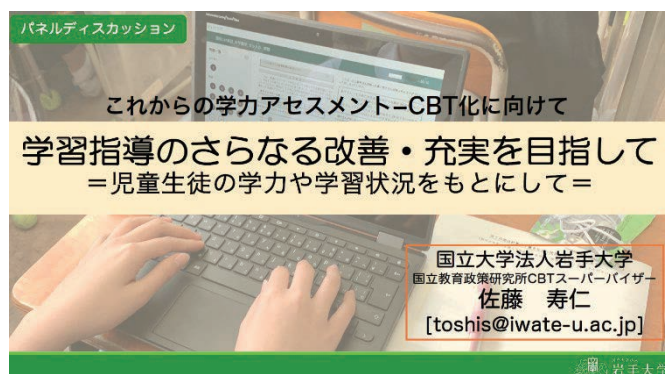
NIER

白水：ありがとうございました。様々なやりたいことがある前の制約をどういうふうに打破していくか。その議論をいただきました。

それでは、佐藤先生、現場の指導改善に、この全国学調、あるいはそのCBT化がどう役立ちそうかの御報告をお願いいたします。

佐藤：はい。岩手大学の佐藤でございます。よろしくお願いします。

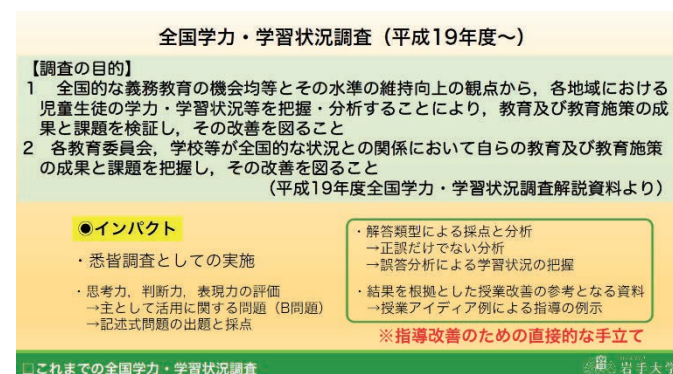
では、画面を共有させていただきます。それでは、よろしくお願いいたします。



私は、平成 28 年から令和 2 年まで、国研の方で、学力調査官をしておりました。実際に PBT ですね、PBT の方で問題作成・分析の方を携わらせていただきました。そういう観点でも見ながらですね、お話をさせていただきたいなというふうに思っております。まず、埼玉県教育局の渡辺課長様よりですね、埼玉県についての取組の御報告をお聞きしました。義務教育数年にわたってですね、児童生徒の学習状況、学

校とかですね、学級とかっていうそういうフォーマルなフレームだけではなくてですね。個々の成長にも着目して、実施されているっていうこと。本当に素晴らしいことだなというふうに思っております。

また、その成長を捉えるだけではなくてですね、今度は学校、先生たちに対しても、そういう結果をフィードバックして、指導の改善に役立てるということで、非常にその自治体としての PDCA のサイクルだから、回っていると充実しているということが伝わっております。あとは、これから CBT 化ということに、務められているということで、いいお話を聞けたなというふうに思います。渡辺課長様、ありがとうございます。私の方からですね。これまでの学力調査を振り返りながらですね、CBT に対しての期待、そのようなものをお話ししていきたいと思えます。



まず、全国調査の取組っていうものは、画面の上にある通りの狙いでですね。19 年度から始まったわけです。当時のインパクトとしてですね、まず、悉皆調査であるということですね。それと当時記述式の問題が出てきましたので、これはかなり学校においてはですね、インパクトがあったのじゃないかな。ただ、単に正誤を書くとかですね、○×をただ書くとかで

ですね、そういうことだけではない解答方法でそういった解答ごとの状況を見てですね、学校の指導に直接的な手立てとしてですね、生かされるような情報がですね、国から下りてくるというようなことが起こったというふうに思います。

全国学力・学習状況調査 (令和元年度～現在)

【調査の目的】
 1 全国的な義務教育の機会均等とその水準の維持向上の観点から、各地域における児童生徒の学力・学習状況等を把握・分析することにより、教育及び教育施策の成果と課題を検証し、その改善を図ること
 2 各教育委員会、学校等が全国的な状況との関係において自らの教育及び教育施策の成果と課題を把握し、その改善を図ること
 (平成19年度全国学力・学習状況調査解説資料より)

調査問題は、新しい学習指導要領が求める育成を目指す資質・能力を踏まえ、それを教育委員会や学校に対して、具体的なメッセージとして示すものとなるよう検討を進める。
 全国的な学力調査の今後の改善方針について(まとめ)
 (平成29年3月)

- ・ 悉皆調査としての実施
- ・ 問題作成の枠組み
 →主として「知識」に関する問題 (A問題)
 →主として「活用」に関する問題 (B問題)
- ・ 解答形式
 短答式、選択式、記述式
- ・ 解説資料、報告書、授業アイデア例

- 新学習指導要領の考え方への各教育委員会や各学校の理解を促すため、従来の「A問題」と「B問題」に区分するといった整理を見直して、一体的に調査問題を構成することとした。
- 各教科においても、枠組みの変更による問題作成方針のアップグレードが行われた。

□ これまでの全国学力・学習状況調査

まりは学習指導要領の理念にかなり寄った形で枠組み自体も変わる。いわゆるアップグレードされたのかなというふうに思っております。

また、学習指導要領との関わりもあるとは思いますが、令和元年度よりですね、少しフレーミングが変わるといことが起こりました。いわゆるA問題B問題、知識活用といった問題を一体的に問うていうことですね。

当時、資質能力の育成ということがありましたので、そこに寄った形です。つまりは学習指導要領の理念にかなり寄った形で枠組み自体も変わる。いわゆるアップグレードされたのかなというふうに思っております。

H19～30年	「活用」の活用構成の枠組み	R01年～	「活用」の活用構成の枠組み
活用型の問題	活用された問題の枠組み(活用)	活用型の問題	活用された問題の枠組み(活用)
活用型の問題	活用された問題の枠組み(活用)	活用型の問題	活用された問題の枠組み(活用)
活用型の問題	活用された問題の枠組み(活用)	活用型の問題	活用された問題の枠組み(活用)

問題発見・解決の過程を重視した問題作成へ

□ これまでの全国学力・学習状況調査

枠組みの変更が起こったことによってですね、各教科でも問題作成の枠組み自体の変更も求められるということが当時起こりました。今画面中にあるのは例えばということで細かくは申し上げませんが、これは中学校の数学の枠組み変更ですね。左側から右側に移ったということになる訳ですけども、学習指導要領の中で問題発見・解決の過程を重視

するということが出てきましたので、問題の中にもそういった文脈とか、かなり制約はあるのですけれども、その中で文脈の中での問題解決ということを意識した問題作成が行われるようになりました。

令和3年度全国学力・学習状況調査大問2 (1)

図1のような直角三角形があります。

式 $3 \times 4 \div 2 = 6$
 答え 6 cm^2 55.2%

(1) 図1の直角三角形の面積は何 cm^2 ですか。
 求める式と答えを書きましょう。

▼水平な辺を底辺としており、三角形の底辺や高さの関係について理解していない…
 ▼三角形の面積を求める公式について理解していない…

▼示された全ての辺の長さに着目して立式しており、必要な情報を選び出すことができていない…

解答類型別の反応率

8.4%	3×4	答え 12 cm^2
5.2%	$5 \times 3 \div 2$	答え 6 cm^2 以外
20.0%	$3 \times 4 \times 5 \div 2$	答え 6 cm^2 以外
5.2%	?	答え 6 cm^2

□ 調査結果の分析を提供する

実際に調査を行う、毎年行っている訳ですので、かなりのデータが出てくる訳です。今画面上に出ているのは算数の問題ですけれども、三角形の面積を求めるといこと。こういう求める際にですね、子供がどういうことに困っているのかということ。正答率だけではなくてですね、誤答からどういうふうを考えているのかなということでは、学校の先生方にはかなり

具体的にですね、問題に寄った形で、あの指導の改善の視点というものを示してきているのかなというふうに思っています。

例えば、こういう三角形の面積を考えるといいですね。底辺がどうだとか、高さがどうだとかって、そういう議論もあると思うのですけれども、報告書にはですね、下の赤字で書いてあるようなことが書かれてあって、例えば3掛ける4掛ける5割る2みたいな何か全部使ってしまったみたいなのが20%ある訳です。こういう子供がどういうふうを考えて、こういう問題解決に直面しているのかということ。分析するにもこういうデータがあることによってですね、進められるということでした。また先生方によってはですね、まあ、そういうものではないかとかですね、前からそんなことを思っていたよってという風におっしゃる方がいます。

いわゆる教師の勘ですね。この教師の勘っていうものがデータというものを合わせることで、より明確なエビデンスになるのではないかと、かなり指標としても、うまくいっているのじゃないかなというふうに思います。

調査結果の分析をもとに作成されている授業アイデア例 ※令和3年度

調査結果の分析を提供する

後は、いろいろな問題はありますが、そういうことを根拠にしながらですね、こういったいわゆる指導をどうやって進めていくのかといったような具体的な例ですね。これはあくまで例なので、こうしなさいってことはもちろん国研は言わないのですけれども、こういった例も示すことで、先生方の指導改善を支えるということをやっております。

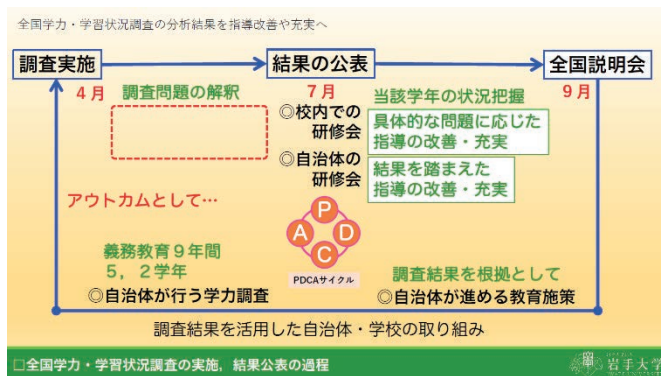
自治体に提供される調査結果を示す資料 ※例：令和4年度岩手県

自治体	生徒数	平均正答率	中央値	標準偏差
岩手県(公立)	8,633	6.8 / 14	49	7.0
全国(公立)	891,913	7.2 / 14	51.4	7.0

○自治体、所管する学校における学習集団としての状況把握
▼学習集団と個の状況の把握 ▼差をどうみるのか

調査結果の分析を提供する

これは各自治体にですね、国研の方から情報提供されるものです。これを見ていただくと分かる通りですね。なかなかこれを見てですね、子供たちの学習状況を緻密に調べていくってのはなかなか厳しいなというふうに僕も思います。なので、ついつい平均正答率だけのポイント差だけ見てですね、本当に差があるのかどうか。統計的には分からないのだけれども、そこだけのポイントだけで学力を見ていくといったようなそういう方向も出てきてしまって、少し望ましくないなというようにも出てきているかと思います。



学力調査っていうのは一つのサイクルの中ですね、取り組まれていると思うのですけれども、調査実施が4月にあって公表が7月にあると、この間がかなり長いと、そして、公表があってからですね、自治体の研修があったり、校内でも研修があったりっていうのはかなりやられているというふうに聞いております。そして、そこから大体、自治体の学力調査が始まったりしてですね、それが埼玉県は少し違う

のですけれども、小学校5年生と中学校2年生、つまり全国調査の前の年ですね、これに学力調査を行う。もちろん、子供たちの状況を形成的に見るという意味ではもちろんいいとは思いますが、少し間違った見方で、学力調査の問題ができればいいみたいな。そういうアウトカムの状態が少し偏ってきているのではないかなというように懸念もあるのではないかと、このように思います。

何のために、誰のためにある全国学力・学習状況調査か

PBTによる全国調査

- 児童生徒一人一人の学習状況の把握と改善
 - ・どんな内容を、どのくらい理解しているか
 - ・記述式による思考・判断・表現の状況把握
 - ・個に応じた手立て
 - ・授業等の指導の改善の根拠となるもの
- 学習集団の状況の把握と改善
 - ・各自治体が所管する学校の状況把握
 - ・教員研修での活用
 - ・教育施策への利用（根拠）

CBTによる全国調査

- 調査実施、結果分析における一連のプロセスの短縮なるか
- 問えなかったことが問えるようになるか
- 正答率などの結果だけでなく、学習者の思考過程に着目すると

→誰のための調査？調査目的？

→調査の枠組みの検討

→用いる統計的手法のメリット・デメリットの確認

□PBTからCBTへ

ただ、これからのPBTからCBTに変わっていくことで考えていかなければならないこと。たくさんの先生がおっしゃっているように、単なるコンピュータベースにするだけではないということなのですね。そうしたときに、いわゆる調査実施と結果分析の間のプロセスですね。この辺の時間的な短縮が本当に可能なのかどうかとか、教科的に言えばですね。紙ではなくてコンピュータにすると、何かこれまで取れなかったものが取れるようになるのではないかと。

つまり、そういったできなかったことができるようになるのではないかとということを経験教育の専門の先生方をどんどん入れてですね、考えなきゃいけないのじゃないかなというふうに思っています。そして、正答率の結果だけではなくて、学習者の思考過程、先ほどのログっていう話もありましたけれども、そういったことから学習者に寄り添ってですね、見ていけるのではないかな。ただ、一方でよく整理をしないといけないなんてこともあると思います。つまり、誰のための調査、何のための調査っていうことですね。そして、これが決まってきたらですね、やはり調査の枠組みっていうものをしっかり検討しなきゃいけないのかなというふうに思っています。今の学力調査はどちらかというと、学習指導要領の理念ですね。そちらの方に向いていると思います。先ほど、石井先生のお話もあったかなというふうに思います。

そして、今回、いわゆるIRTといったような統計的手法、これのメリット・デメリットというのは必ずあると思います。このメリットをですね、学校にどう伝えていくのかということもあるのではないかなというふうに思っています。実際に指導されるのは先生方なので、この辺をどう伝えていくということがポイントかなというふうに思っております。

何のために、誰のためにある全国学力・学習状況調査か

PBTによる全国調査

EBPM 国

EBPM 自治体

PDCAサイクル

学校（教師）

PDCAサイクル

教科教育教育方法 など

地域の子ども達一人一人の資質・能力を育成したい
学校づくりのための一つの観点として

CBTによる全国調査

- 調査実施、結果分析といった一連のプロセスの短縮なるか
- 問えなかったことが問えるようになるか
- 正答率などの結果だけでなく、学習者の思考過程がわかるか

→誰のための調査？調査目的？

→調査の枠組みの検討

→用いる統計的手法のメリット・デメリットの確認

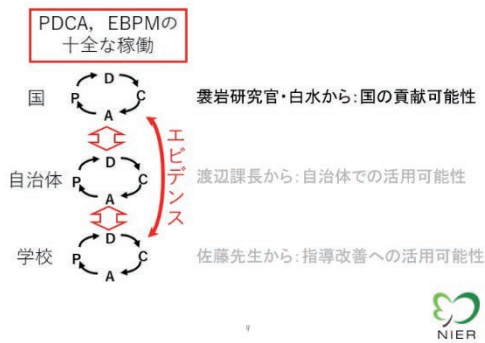
□PBTからCBTへ

CBTになっていくことによってですね、国・自治体・学校・教師ですね、いろいろなところでのPDCAが回ると思うのですけれども、いろいろな専門的な方に入っていくただきながら、常にブラッシュをしていくことは大事じゃないかな。その中に教科教育とかですね、専門家も入れながらですね、良い方向に改善で回していくということが大事なのではないかなというふうに

思いました。少し早口になりましたけれど、私の方から話を終わりたいと思います。以上です。

白水：佐藤先生ありがとうございました。3つのサイクルを結び付けて回していけるとよいのだけれども、そのためにはどんな設計が必要だろうかという御提言を頂きました。

それでは、最後、私たちの方から国は一体何ができるのか考えてみたいと思います。



様々なご示唆, 要望

「毎年、国と私たちの自治体の平均、自治体と学校の平均を見比べて課題や成果を見つけてきたけど、これだけでいいのかな」

「問題一つひとつの児童生徒のつまずきやできはわかるのだけど、それがどういう資質・能力の育成を表しているのかしら」

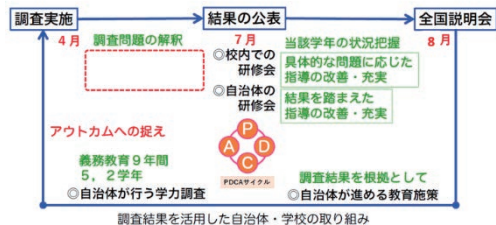
「自治体の問題づくりは限界があるから、国のものを共有できないかな」

「教育政策が先生方の組織的な授業改善や児童生徒の学力向上につながっているだろうか」



限界があるので、国のものを共有できないか。あるいは教育政策が先生方の組織的な授業改善や児童生徒の学力向上につながっているのかどうか。このような疑問を頂くこともございます。

佐藤先生スライドより



これを先ほどの佐藤先生の枠組みに当てはめてみますと、上の方の調査をやって、実際に問題をそれぞれの子供たちがどう解いたかという世界と、下の方の実際にそれがあつた種の点数・値になって、結果として返ってきたときに、それを教育施策としてどう進めていくか、さらにはアウトカムといったときに、この一つ一つの問題から集まった点数というのが、子供たちの資質・能力の一体どういう発揮、育成を表しているか、この両者を結び付けていくのがなかなか難しいと

いうことではないか。では、これはもう別の物でやってしまおうということになるのか。

様々なご示唆, 要望

「毎年、国と私たちの自治体の平均、自治体と学校の平均を見比べて課題や成果を見つけてきたけど、これだけでいいのかな」

「問題一つひとつの児童生徒のつまずきやできはわかるのだけど、それがどういう資質・能力

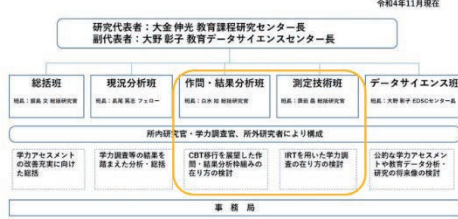
3つのサイクルを連携させながら、これらの問いにしっかり答えを出していくために、国研にできること、やるべきこと

「教育政策が先生方の組織的な授業改善や児童生徒の学力向上につながっているだろうか」



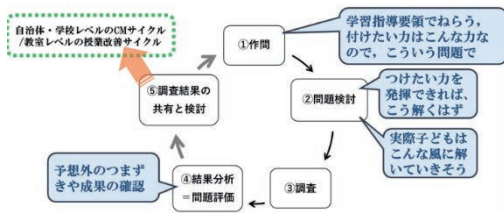
それとも、先ほどあつたようなこういう問いに対して、正に3つのサイクルを連携させながら、しっかり答えを出していけるのだろうか。そのために、文科省の下支えとなるような基礎資料を国研の方から提供していただけるだろうか。そういうことを考えていくというのが、この学力プロジェクトの大きな目的ではないか。そんなふうに考えます。

「学力アセスメントの在り方に関する調査研究」の研究体制



今日は作問・結果分析班の班長を務めております私と測定技術班の巖先生が来ていただいておりますので、このそれぞれの取組はほんの一部ですけれども、御紹介していきます。

学びの質を支える学力調査のPDCAサイクル確立を目指して



作問・結果分析班の方としては、まず作問・結果分析のPDCAのプロセスサイクルというのを充実していくということを目指しました。

問題を作ったら、それが果たしてひとつひとつの問題が、学習指導要領が狙うどのような資質・能力とアライン（結び付いて）しているのか。どういう力のために、この問題を出していきたいのか。その問題を実際出してみたときにその力を発揮できる子供であれば、それはどういうふうに分かれるのか、そのプロセスですね、

それを教科の観点だけではなくて、下にあるように、子供はどのように解くかということを考えながら、検討できるとよいのではないかとことです。

佐藤先生スライドより

数学的プロセス	数と式	図形	関数	資料の活用
<p>算数・算理・算術に迫る場面</p> <p>算数における問題</p> <p>算数・算理・算術に迫る場面</p> <p>算数における問題</p> <p>算数・算理・算術に迫る場面</p>	<p>算数・算理・算術に迫る場面</p> <p>算数における問題</p> <p>算数・算理・算術に迫る場面</p> <p>算数における問題</p> <p>算数・算理・算術に迫る場面</p>	<p>算数・算理・算術に迫る場面</p> <p>算数における問題</p> <p>算数・算理・算術に迫る場面</p> <p>算数における問題</p> <p>算数・算理・算術に迫る場面</p>	<p>算数・算理・算術に迫る場面</p> <p>算数における問題</p> <p>算数・算理・算術に迫る場面</p> <p>算数における問題</p> <p>算数・算理・算術に迫る場面</p>	<p>算数・算理・算術に迫る場面</p> <p>算数における問題</p> <p>算数・算理・算術に迫る場面</p> <p>算数における問題</p> <p>算数・算理・算術に迫る場面</p>

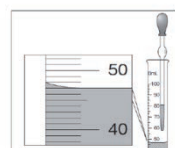
先ほど、佐藤先生のスライドにあったように、今回の令和元年度からの枠組みにおいては数学的なプロセスというのが出てきました。いろいろな単元の中身であっても、数学的なプロセスで問題を解いている子はどんなふうに分かれるのだろう。それを予想しながら、問題を作ってみる。そうすると、それがあつた種の仮説になりますので、実際に結果が出てきたときに、予想外のつまずきやでき方が見えてくる。テスト理論や教育測定による統計的な結果の見取りというのでも深まって、このサイクル

を、まずはしっかり良問を作るために回していくというのが自治体、学校レベル、そして教室レベルでの指導改善のサイクルの充実につながっていくのではないかとくろんでおります。

本年度小学校理科調査 結果報告書より

2(1)(2) 器具の名称の理解・正しい扱い方 知識・技能

図 (1) の器具に、次の図のように、50 の目盛りよりも下まで水を入れました。50 mL の水をはかりきるためには、このあとスポイトでどれだけの水を入れるとよいですか。下の 1 から 4 までの中から 1 つ選んで、その番号を書きなさい。



どの誤答が一番多そうでしょうか？
(正答2 70.2% 誤答1 17.3%)
どんな学力層の児童がその誤答をしそうでしょうか？



では、この充実を図るために、今のPBTの状態、しかもIRTが入っていない状態でどのようなことを始めているかを少し紹介していきます。

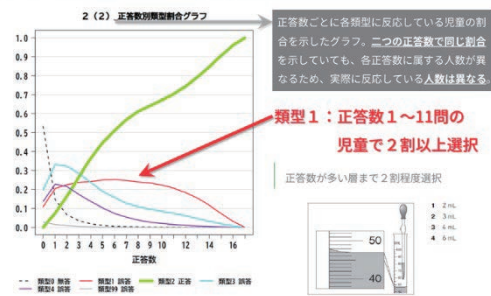
これは、今年小学校理科の調査で出された問題です。メスシリンダーで測るときに、50ミリリットル測りたいのだけれども、あと何ミリリットル入ると50ミリリットルになるか。

その問題に対して正答は3ミリリットルですが、誤答の2ミリリットル(1番)というのを選んで

いる子供たちが17.3%います。ただ、この結果というのは、まとまった結果ですので、どんな学

力層の児童がその誤答をしたかというのはいけません。

2(2) 正答数別類型割合による分析



これに対して、今年度の結果説明会の資料に取り入れられたのが、例えばこのような資料でございました。

見方としては、横軸が学力を表すと思ってください。どうやって学力を見ているかという、その年に出された17問中何問あっているか、です。17問全部あった子もいれば、1問もあわない子もいるというのを、取りあえず学力の高い低いと見て、その児童がこのメスシリンダーの問題に対してどういう解答したかということを示しております。緑が正答を選んだ子供たちで、17問

全体の正答率が上がれば、この問題も正答できるようになっています。注目したいのは、この【類型1 誤答】を2割ぐらいいらっしゃる児童というのが結構たくさんいて、全部で11問あっているような子でも2割ぐらいいらっしゃるということの意味しています。

2(2) 授業改善

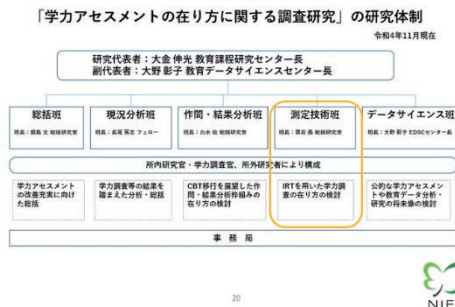
目的に応じて器具や機器を選択し、
観察、実験などに関する基本的な技能を
身に付けることができるようにする

50mLの水を正確にはかり取るという目的に合うメスシリンダーを選択するために、
その役割や目盛りの読み方を確認した上で、
はかり取りたい水の量より少ない水の量を最初に入れる理由や、
足りない分の水を入れる際、スポイトの先が水の中に入らないようにする理由について話し合い、
確かめ合う学習活動が考えられる。

全体の正答数からすると、これは単に「知らなかったから」という原因かもしれないというので、解説の方には、「目的に応じて器具や機器を選択し、観察、実験などに関する基本的な技能を身につける」とよいのではないかと書かれております。

このように結果が詳しく見えるようになってくると、単に全体で何問できたかという結果、あるいは逆にこの問題だけが全員でどれくらいできたかという話ではなくて、問題ごとにこういう層の子供たちはこんなふうに取り組むのだな、というそれぞれの特徴というのが見えてきます。これを発展的にやるようなものがIRTだと考えてみると、要点としては、子供が問題を解いて、何ができているか、何につまずいているか、それを更に詳しく知りたい、というのが、IRT・CBT導入の背景にあるのではないかと思います。

それでは測定技術班の巖先生の方から、その取組の一部を御紹介いただければと思います。



全国学力・学習状況調査への 項目反応理論の導入

- 現行の悉皆調査
 - 全国同日一斉実施
 - 調査問題は1セットに限定
 - 全員が同じ日に同じ問題セットを解く
 - 出題される問題数及び領域が限られている
- CBT移行後の悉皆調査
 - 複数日に分散実施
 - 調査問題は複数セット用意
 - 別の日に受ける/異なる問題セットを解く
 - 幅広い領域からの出題



ます。大津先生の発表でありました現行の悉皆調査は、一斉実施、ワンセットで行っています。CBT移行後は、分散実施、複数日で複数のセットで行う。こういうことが検討されておりまして、何でこれをやると項目反応理論というものが必要になるか。今までの皆さんの議論の中に、実は項目反応理論、IRT というのが出てきていたのですが、これ、なぜ必要なかっていうことを少し補足させていただくと、問題セットが異なっても能力を比較可能にする。

巖岩：測定技術班の巖岩と申します。よろしく申し上げます。測定技術班は、項目反応理論を全国学力・学習状況調査にどう導入する、導入の仕方を検討している班になっております。

具体的にはいろいろなことをやっているのですが、実際に使われたときにならないと多分皆様の目に触れることもなく、しかも得点しか出てこないで裏方に回る事柄なのですが、補足的なところを述べさせていただきます。

全国学力・学習状況調査への 項目反応理論の導入

- なぜ項目反応理論が必要なのか
 - 問題セットが異なっても能力を比較可能にするため



- 問題が異なるセット間では、正答数や正答率を用いて能力を比較することができない



このために、使われているのがIRTでございます。

これ全く架空の例なのですが、2つのセット。例えばある人がセット1に答え、ある人がセット2に答える。問題が異なるセットを使っている場合、まず当然のことながら、問題が違うので難しさが違います。正答率っていうのは比較できないのです。そこで項目反応理論を使うのですが、項目反応理論とは何か。

全国学力・学習状況調査への 項目反応理論の導入

- 項目反応理論とは何か
 - ある問題 (=項目) に正答するのか、誤答するのか (=反応)、その確率を「問題の特性 (難易度等)」と「受検者の能力」の関数として表す
- ↓この関数を使って
- 受検者の問題に対する正誤データ (項目反応) から「問題の特性 (難易度等)」と「受検者の能力」を推定する

※経年調査では、項目反応理論を利用した経年変化の把握が可能になっている

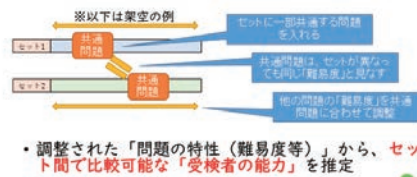


か、受検者の能力を推定するということを行います。これが大津先生の発表の中で出ていました経年調査では経年変化の把握が可能になっているという根拠なのです。

一言で言うところある問題、ずっと項目とかいう言葉が出ていましたが、ある問題のことを項目と呼びます。正答するのか誤答するのか、これを反応と呼んでいます。その確率、その反応が起こる確率を問題の特性、難しさとかですね、と、受検者を能力の関数として表すっていうのが項目反応理論なのですが、ただそれだけだとよく分からないと思うのです。この関数を使って、受検者の問題に対する、正答・誤答データ、これは調査データなのですが、この項目反応を表すデータから問題の特性と

全国学力・学習状況調査への 項目反応理論の導入

- なぜ項目反応理論が必要なのか
- 問題セットが異なっても能力を比較可能にするため



能力を推定する。こういうことをやろうと、測定技術班では、まずどういうやり方がいいのかわかっていうのをこういういろいろ検討しております、こういう経年比較とか複数セットの利用、そしてPISAだったら適応型テストとか様々なメリットがあるのですが、これはあくまでも道具にすぎないので、その生かし方っていうのは測定技術班を超えて、我々のプロジェクト研究、又は、文科省の中でワーキンググループが、今検討しているという状況でございます。私からは以上でございます。

教育の未来をみんなで創るために 共に答えを探す力をお貸しください

「毎年、国と私たちの自治体の平均、自治体と学校の平均を見比べて課題や成果を見つけてきたけど、これだけでいいのかな」

「問題一つひとつの児童生徒のつまずきやできはわかるのだけど、それがどういう資質・能力の育成を表しているのかしら」

「自治体の問題づくりは限界があるから、国のものを共有できないかな」

「教育政策が先生方の組織的な授業改善や児童生徒の学力向上につながっているだろうか」



白水：ありがとうございました。最後のこのスライドにあるような問題に対して答えを出していくというのは、簡単なことではない。それを皆さんの力を借りながら進んでいきたいと思えます。それでは、池田先生、佐藤先生、渡辺課長の順に一言ずつ、1分ほどでコメント・御質問、お願いいたします。

池田：ありがとうございます。能力という見えないものをテストっていう見える化して、それを測っている。その結果で政策に反映していくっていう、その全体像がとてもよく分かりました。ありがとうございました。一つ提案です。問題を作る方は、どういう能力を測っているかというのをすごく考えながら作っているっていうのはよく分かったのですが、それをもっと見えるようにしていくっていうのが多分学校の先生にとってとても助けになるのではないかなと。まずフレームワーク・枠組みがあって、問題が何を測っているのかというのをクリアにし、もし、つまずきがあれば、どういうつまずきをした子はどういう能力が欠けているのかっていうのをクリアにしていくことによって、先生方は日々の実践に活用できるのではないかと思います。

そして、コンピュータ化することによって、そうした情報が提供しやすくなる。公開されている問題ではなくて、類似問題も作って、それをフレームワークに当てはめることによって先生方が使える問題を増やしていく。そして、それぞれ問題が何を測っているのかっていうのを詳しく説明する文章をネットで配信することによって、先生方が10分とか20分とかの時間を使って生徒にテストを受けてもらって、その結果を元に実践に反映するというようなこともできていくのではないかと思います。

白水：はい。すばらしい御指摘、ありがとうございます。佐藤先生、お願いします。

佐藤：はい。では、二つのことを話します。一つはですね。先ほど、白水先生がおっしゃって

ただいたように、今までの学力調査は一つの問題に対してどういう反応しているかということだけで分析されてきているのですよね。しかし、先ほどの理科の例を取り上げていただくように、背後にあるものをどうつかんでいくのか。そういう子供たちの本当の意味での特性っていうものをどうやって測っていくのかっていうのがこれに期待したいなって、思っております。

それと、もう一つ気になるのは能力という言葉が、僕はいつも気になっていて、何かこう学習指導要領に慣れているので、能力って言うと資質・能力かなってすぐ言うてしまうのですけれど、そもそもでは能力って何なのかなと。今、池田先生もおっしゃっていましたが、これを学校にどういうふうに説明しながら、そして学校の先生がそれをどういうふうに平たくやっていくのかなとかって考えると、もう少し説明が欲しいかなというふうには思っております。以上です。

白水：ありがとうございました。それでは、渡辺課長、最後お願いします。

渡辺：はい。CBT・IRTによって、その可能性は無限大になった分、そのデータをいかに使うか、あるいはどう使うと効果的なのかが、これからの課題と考えています。やはり、どうしても学校の現場あるいは県単体では、そういった部分の検討は限界があるので、是非国立教育政策研究所には、好事例を示したり、その道しるべとなっただけならば、我々としては非常に有り難いので、引き続きお願いできればと思います。

白水：ありがとうございました。重要なコメントを頂きました。

教育の未来をみんなで創るために 共に答えを探す力をお貸しください

「毎年、国と私たちの自治体の平均、自治体と学校の平均を見比べて課題や成果を見つけてきたけど、これだけでいいのかな」

「問題一つひとつの児童生徒のつまずきやできはわかるのだけど、それがどういう資質・能力の育成を表しているのかしら」

「自治体の問題づくりは限界があるから、国のものを共有できないかな」

「教育政策が先生方の組織的な授業改善や児童生徒の学力向上につながっているだろうか」



最後のスライドに対して、今、現時点で考えられるお答えを簡単に申し上げますと、最初の「点（正答率）」というものが資質・能力の表れであるべきであり、逆に一つ一つの問題が資質・能力の発揮の場になっていて、だからこそ、問題だけを国と自治体間で共有するのではなくて、その問題の作り方、結果の読み取り方のサイクルそのものを共有していくことができるとよいのではないかと。それによって、現場に近い自治体と知見を共有しているからこそ、時々立ち止まって、「本当にテストだけで大丈夫なのだろうか?」、「日々現場で発揮してもらっている、身につけている子供たちの資質・能力は他にもあるのではないだろうか?」、「それをテスト以外のどんな観察の窓で見られるのだろうか?」ということを考え直せるのではないかと。こうした答えを探す取組を、各現場の先生方、専門家の先生方など、みなさまの力を借りながら進めていければと考えております。これで終わりにさせていただきます。

どうもありがとうございました。

閉会挨拶

大野 彰子：国立教育政策研究所教育データサイエンスセンター長

大野：本日は参加の皆様、誠にありがとうございました。本日、約 530 名の御参加を頂いております。講演者、パネリストを始め、皆様の御協力によりまして予定したプログラムを進めることができました。厚くお礼を申し上げます。

学力アセスメントの在り方というのは、時代や社会の変化とともに変わっていくものだと思います。

本日、その学力アセスメントというのは、何を見たいのかという目的に合わせた設計が重要であることが強調されたところですし、それから学力アセスメントで得られたそのデータをどう活用していくかというところには、様々な発展の可能性があると思います。本日、そのような広がりについても皆様の理解が一層進みまして、それぞれのお立場での今後の取組などに生かされる機会ともなればと思っております。

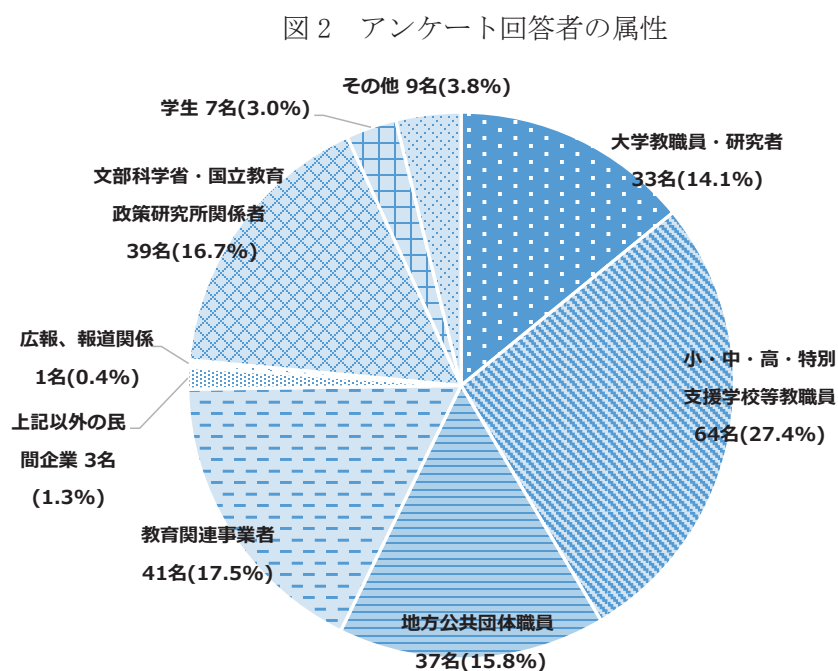
時間の関係から Q&A の時間を設けることができませんでしたので、是非この後案内があります、アンケートに皆様からの御意見、御感想をいただければと思っております。今後も本プロジェクト及び研究所の取組全般に関しまして、御理解御支援を賜りますようお願い申し上げます、簡単ではございますが、閉会に当たっての御挨拶といたします。本日は御参加誠にありがとうございました。

第3章 アンケート結果から得られた示唆

以下では、本シンポジウム終了後に行ったウェブアンケートの結果を取り上げる。紙面の都合上、アンケートの詳細は割愛し、主要な結果のみを掲載する。なお、本報告書にて公表した結果は、公表することに同意いただいた参加者のものである。

参加者のアンケート回答者数は 234 件であった。これは本シンポジウムの参加申込者数 768 名の 30.5%、当日の参加者数 562 名の 41.6%に相当する。

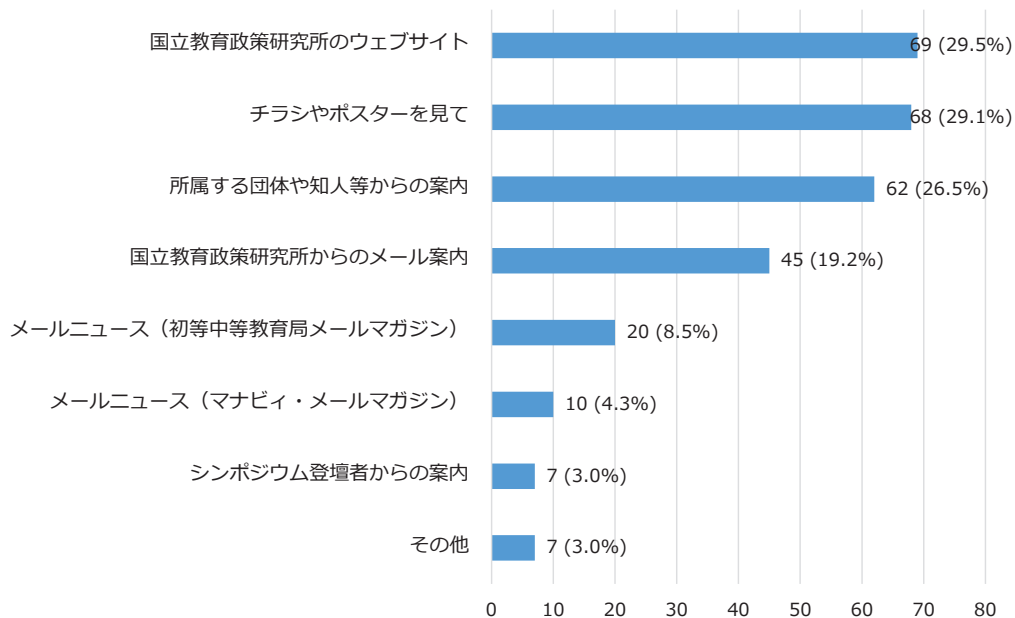
アンケート回答者の属性について、内訳を図 2 に示す。



属性として最も多かったのは、小・中・高・特別支援学校等教職員 64 名 (27.4%)、次いで、教育関連事業者 41 名 (17.5%)、文部科学省・国立教育政策研究所関係者 39 名 (16.7%) であった。GIGA スクール構想の推進や教育の情報化の拡大を背景として、教職員や教育関連の事業者、国や地方の職員など、多様な属性の方に興味を持っていただけたのではないかと考えられる。

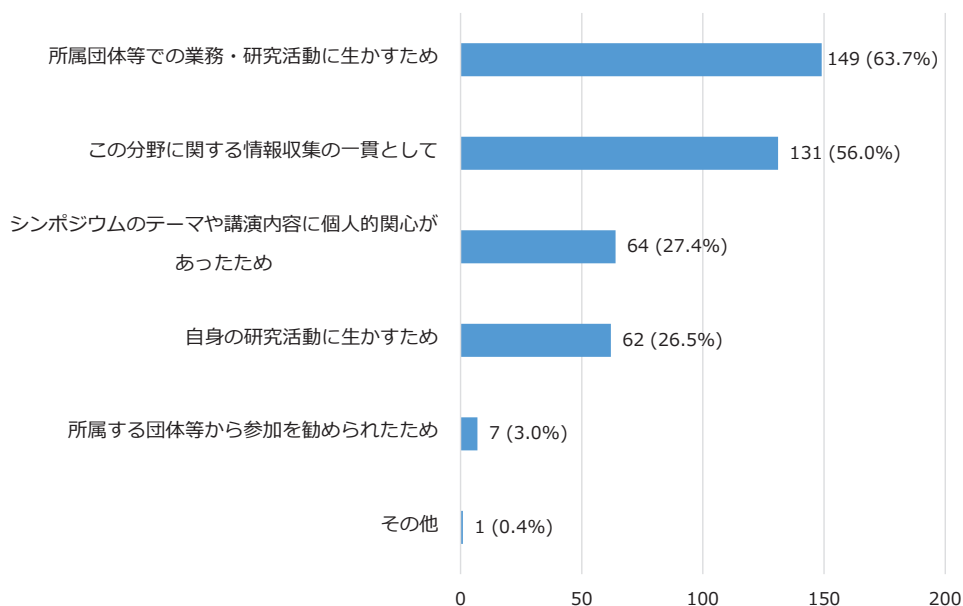
本シンポジウムを知ったきっかけ（複数回答可）としては、「国立教育政策研究所のウェブサイト」が69名（29.5%）、「チラシやポスターを見て」が68名（29.1%）、「所属する団体や知人等からの案内」が62名（26.5%）、「国立教育政策研究所からのメール案内」が45名（19.2%）と多様であった（詳細は図3に記載）。

図3 本シンポジウムを知ったきっかけ



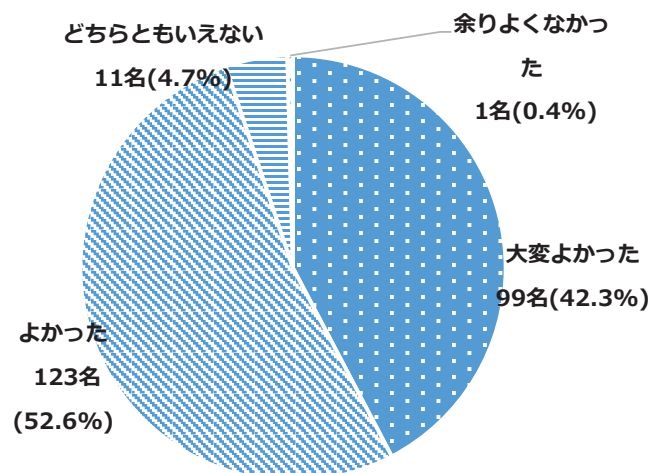
本シンポジウムに参加した目的（複数回答可）については、「所属団体等での業務・研究活動に生かすため」が149名（63.7%）、「この分野に関する情報収集の一貫として」が131名（56.0%）、「シンポジウムのテーマや講演内容に個人的関心があったため」が64名（27.4%）であった（詳細は図4に記載）。

図4 本シンポジウムに参加した目的



シンポジウム全体の満足度は、「大変よかった」と「よかった」を合わせて94.9%となった(図5)。個別では、「大変参考になった」と「ある程度参考になった」を合わせて基調講演が97.3%、講演①が96.9%、講演②が92.3%、パネルディスカッションが97.7%となっており、ほとんどの参加者にとって満足度が高かった。

図5 シンポジウム全体の満足度



シンポジウム全体に関する感想・意見について、全体で120件の回答が寄せられた。全体の傾向として、大変勉強になった、発表が丁寧でわかりやすかった、CBT化の動向がよく分かった、教育データ利活用の将来可能性に期待する、といった回答などが多く寄せられた。

以下では、意見や感想の内容ごとに分けて一部を列挙する(なお、個人が特定されないよう、また表記を整える観点から、表現を適宜編集している)。

◇全体について：CBTに関する動向把握

まず、全国学力・学習状況調査のCBT化や海外の動向、CBTのメリット・デメリットに関する情報を把握することができ、有意義であったという感想を多く頂いた(意見1~6など)。

1. 学力アセスメントについて学びたく、シンポジウムに参加させていただきました。どのような展望をもってCBT化を進めていくのか理解ができ、心構えをもつことにつながりました。
2. 今後、全国学力・学習状況調査がCBT化される上で、大変参考となる情報を聞くことができた。メリットとデメリットを踏まえた上で対応していきたい。

3. 想像していた以上に、CBT化が着実に、そして急速に進んでいることがよくわかりました。また、その意義や課題についても大変勉強になりました。ありがとうございました。
4. CBTの導入は教育関係者にとって非常に関心の高いテーマだと思います。OECDの取組から埼玉県での取組まで、幅広い関係者で様々な観点から議論されていたのがよかったと思います。
5. 今後の文科省の方針、国際動向に関する貴重な情報を得ることができ、今後のCBT開発方針の設計の参考となった。
6. 学力調査に関わる動向について、海外の状況と比較しながら、今後の在り方を考えることができたように思います。CBT化を前提に進めることは時代の流れだと思ふ反面、不易で大切にしなければならない部分もあるはずなので、そのバランスをいかに担保していくかが今後の課題だと感じました。

◇全体について：教育データの利活用

教育データの利活用についても、下記の意見が寄せられた。具体的には、教育現場での活用を考えていきたい、といった意見がみられた（意見7～10など）。また、意見11～13では、項目反応理論の活用や展望を含め、教育データ分析の必要性を感じたという感想がみられた。

7. 大変有意義な時間になりました。未来ある子供たちのために有効なデータ駆動化を目指したいですね。
8. 様々な方が簡潔に考えを話してくださり、話の内容はとても分かりやすかったです。本日の内容を、実際の学校現場における学習指導（を含めた全ての教育活動）でどのようにいかせるか考えていきたいと思いました。学校現場では、多くのデータを読み込んだり、分析したりすることは不可能に近いのですが、目には見えない個々の学力について、蓄積されるデータをアセスメントにうまくいかせるようになりたいと思います。
9. テストの○×の結果だけでなく、そのプロセスや生徒の層について分析できること、そこから学力を上げるための課題が発見できることに感銘を受けました。ありがとうございました。
10. 大変勉強になりました。CBTで集めたデータをどのように分析・活用できるのか、より掘り下げた内容をお聞きしたいので、シンポジウムの第2弾を開催いただきたいと思いました。
11. 各種学力調査の回数を重ねていくことで、多くのデータが収集され、分析された結果をもとに、今後どのように取り組んでいけばよいのか考えるようになってきました。一方で、取り組んできたことが本当に果たして成果が得られるものかどうか判断ができずにいました。本日の講演を聴いて、どの項目とどの項目に相関があるのか具体的に分析していくことの必要性を認識し、子供たちのために、今後どのような方策を講じていけばよいか更に考えていきたいと思っています。ありがとうございました。

12. 様々な教育政策に EBPM が求められている昨今、項目反応理論は生徒の能力を加味しながら公平性を保つ理論ということで、今後の活用事例に非常に興味がありますし、活用の必要性も改めて感じました。ありがとうございました。
13. 全国学力・学習状況調査について、よくわかった。これまでも分析を行い、指導にいかしてきたが、シンポジウムに参加し、これまでの分析の仕方を改善していく必要があると感じた。さらに、全国学調の意義もよくわかり、これまで以上に前向きに取り組めると感じた。

◇全体について：CBT 化に向けた課題

CBT 化に向けた課題として、CBT 化の目的を明確にする必要性や、自治体や学校現場での負担感の考慮などがあげられた（意見 14～17 など）。

14. CBT になると、これまでとは異質の問題も出題できるようになりますが、お話を通して、どういう資質・能力を測るための問題なのかという目的を見失わないようにする必要があると思いました。
15. 学力アセスメント（テスト）の目的をはっきりさせる必要がある、ことを学びました。現場の指導改善が目的なのか、学校の説明責任を果たすことが目的なのか、政策のためのエビデンスを得ることが目的なのか。その上で、データサイエンスの観点からどのように分析するのかについてもっと議論が聞きたかったです。
16. CBT 化の話については、改めて目的レベルの話と方法（やその現実的課題）レベルの話をいったん分けて考えた上で、目的を明確にした上で（その目的の妥当性を十分に議論した上で）現時点でリスクをとっても取組を進めるべき点とそうでない点を整理していくような議論の進め方ができるとよいと感じました。
17. 全国学力・学習状況調査の CBT 化の将来的な方向性を知ることが出来たのは良かった。正直、令和 6 年度実施は、もろもろの課題を解決し、実施するには時間が足りないように思う（世界的に日本が遅れているから、急ぎたいのはわかるが）。講演を聴いていて、PBT と比較しての優位性はそれほど感じなかった。専門家の考えだけでなく、自治体や学校現場での負担感を考慮し、何より児童生徒に有効な手立てとなるような研究と実践になるよう、切に願う。今日は貴重なお話をありがとうございました。

◇全体について：その他

その他の意見として、「理解するのが難しい用語があった」、「どちらかという状況の整理・共有が主眼となっていた」、「もう少し、CBT の具体（背景理論や技術要素等）やデータ利活用の最先端動向に関する話を期待していた」といった御意見が寄せられた。

◇各プログラムについて

各プログラムについて、具体的な意見や感想を頂いた。基調講演については、内容がわかりやすかった、参考になったといった感想がみられた（意見 18 など）。講演①、講演②については、大変勉強になった、参考になったという感想や（意見 19、20 など）、多面的に検討する必要があるという意見もみられた（意見 20 など）。パネルディスカッションについて、非常に参考になった、分かりやすかったといった感想が寄せられた（意見 21 など）。

18. 基調講演の内容がわかりやすかった。

19. 池田氏の御講演が非常に明せきで、大変勉強になりました。

20. 「米国の公的な学力アセスメントについて」の講座で、アメリカの目的別学力テストの具体を知り、目的を複合的に（欲張って）実施される学力テストについては、誤解を招きやすいことを改めて認識できた。CBTによるアセスメントにより、よりきめ細かく、児童生徒の実態に寄り添う指導方針が立てられるようになることは有効だと思う反面、「学力が高いとは何ができることか」「ビックデータをどのように利活用するのか」という点において、暴走する怖さを感じ、慎重に多角的・多面的に協議しながら「よりよいもの」を構築していく必要を感じた。

21. パネルディスカッションの内容は非常に参考になりました。モデレーターの白水先生のすすめ方もよかったですし、佐藤先生、巖倉先生のお話も具体的でよかったです。特に巖倉先生の項目反応理論の御説明はとてもわかりやすく、理解しやすかったです。ありがとうございました。

◇運営について

パネルディスカッション、及びシンポジウム全体の時間配分について、より時間をとってもらいたかったといった意見がみられた（意見 22、23 など）。

22. パネルディスカッションでの各所の提案がレベルが高く、大変興味深く聞くことができました。時間の都合上仕方がなかったことだと思いますが、ディスカッションの時間がもっと長いと多様な意見が聞けたのではないかと思います。

23. パネルディスカッションの時間がもう少し増えて、CBT 化の課題の深堀や解消方法の模索、CBT 化の未来の可能性などが語り合えると、更に有益な時間になったと思いました。

そのほかの運営面における要望として、見逃し配信の希望がいくつか寄せられた（意見 24 など）。これについては、本シンポジウムの動画を公開し、周知している。アーカイブ動画の公開は、今後のシンポジウム等企画においても引き続き行う必要があると考える。また、事前音声確認用の BGM の要望、事前予習の希望、オンライン開催に対する肯定的な意見などがみられた（意見 25 ～28 など）。

24. 「見逃し配信」をしていただけるとうれしいです。（ところどころ抜けざるを得なかったので、もう一度見たいです。）
25. オンライン配信の環境を確認するために、開始前にテスト用 BGM を流しておき、参加者自身にて確認できる等あれば良かったのかも知れません。
26. 一部、IRT についての説明もありましたが、CBT を議論する上で必要な知識については、視聴者に事前に予習しておいてもらおうと、議論への理解が深まるように思いました。
27. zoom でのシンポジウムは、職場にいながら参加でき、公務への負担も少なく、かつ最新の情報に触れることができるので大変有意義と感じている。
28. 地方にいてもこうして国や先進地域の動向を得られるのは大変有り難い。

最後に、国立教育政策研究所に期待する「イベント企画や情報提供」に関する要望や提案として、59 件が寄せられた。これらを「CBT に関する要望」、「データサイエンスに関する要望」、「その他具体的な要望」、「抽象的な要望」、「内容以外の要望」、「その他」に分類した。その結果、「CBT に関する要望」が全体の 28.8%、「データサイエンスに関する要望」が 11.9%、「その他具体的な要望」が 25.4%、「抽象的な要望」が 18.6%「内容以外の要望」が 3.4%、「その他」が 11.9% となった。

「CBT に関する要望」としては、CBT 化に向けての環境構築や問題例、全国学力・学習状況調査における CBT について話を聞けるとよいという提案がみられた。また、CBT の導入による学びの履歴など、可視化された学校教育の未来がイメージできるような話を聞きたいという要望もあった。「データサイエンスに関する要望」としては、スタディログの活用例、全国学力・学習状況調査などにおけるデータ分析の手法や結果、データサイエンス教育など、データ利活用に関するイベントの提案がみられた。「その他具体的な要望」としては、問題作成、教育の情報化、学力や ICT 環境の格差、ICT 活用とアクティブラーニングの有用な実践・実績など様々な観点から提案があった。「抽象的な要望」としては、効果的な実践紹介や教育施策などについて話を聞きたいという要望がみられた。また、「国研は、教育行政と現場の教育（教師）のパイプ役としてとても重要な機能を果たしていると思います。これからの教育の在り方を展望するためにホットな話題を提供していただくことはもちろんのこと、双方向の情報交換ができればいいなと感じています。」といっ

た国研の役割に言及した意見もみられた。また、「内容以外の要望」として、ディスカッションの要素への期待について言及があった。

令和4年度 教育研究公開シンポジウム報告書

学力アセスメントの動向と展望
～ CBT 化に向けて ～

(プロジェクト研究「学力アセスメントの在り方に関する調査研究」)

令和5年(2023年)3月

発行所 国立教育政策研究所

住 所 〒100-8951

東京都千代田区霞が関3丁目2番2号

印 刷 株式会社ワーナー