

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

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

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「論点整理」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することが望ましい）→ 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

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

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

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）

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が）
 - 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

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）

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

受験者 1

問題A

問題B

異なる受験者に共通問題を実施

受験者 2

問題A

問題C

受験者 3

問題B

問題C

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

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、ストーリーを作りだ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

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。