

## 認知過程の想定と検証を軸とした作問・分析プロセスの確立に向けて

Establishing a High-Quality PDCA Cycle for National Assessment of Academic Ability  
by Focusing on Assumption and Verification of Student Cognitive Processes

齊藤 萌木<sup>\*1</sup>、白水 始<sup>\*2</sup>

SAITO Moegi and SHIROUZU Hajime

### Abstract

This paper discusses a research project aimed at establishing a high-quality PDCA cycle for the National Assessment of Academic Ability. A key focus of the project was the assumption and verification of students' cognitive processes within the PDCA framework. This study utilizes the research framework of the learning sciences, which integrates diverse academic fields and practitioners' wisdom. The study, conducted as a workshop, involves data analysis experts and practitioners responsible for organizing surveys. The findings indicate that concentrating on cognitive processes allows for a deeper understanding of actual learning dynamics beyond test answers. This focus is beneficial for improving assessment tasks and developing higher-quality feedback materials that reflect the patterns and diversity of students' thinking processes. Additionally, the study highlights a future challenge: integrating PDCA-based assessment improvements at the national level with curriculum management at the local government, school, and teacher levels.

\*1 聖心女子大学教育学科 専任講師

\*2 初等中等教育研究部 副部長／教育データサイエンスセンター 副センター長

# 1. はじめに

国内外の多くのアセスメントで、その重点が児童・生徒の学習成果の評定・序列化機能から、学習過程を把握・改善する形成的評価機能に移りつつある (Scardamalia et al., 2012)。全国学力・学習状況調査 (以下「全国学調」と略す) も、その目的の一つは「児童生徒への教育指導の充実や学習状況の改善等に役立てる」ことにあり、そのために「教育に関する継続的な検証改善サイクルを確立する」ことが目指されている。

こうしたアセスメントに IRT (Item Response Theory: 項目反応理論) が導入されているのは、能力のより正確な推定に加え、経年変化を追いやすくするなど「伸び」の評価も期待されているためである。一方で、こうした評価を精緻化する動向は図 1 a で示したように、アセスメントを担う関係者間の断絶を生みかねない。すなわち、正答率や素点といった素朴な指標から「能力値」という解釈に素養がいる指標への変更や秘匿される問題の存在など、統計的な側面が強調されることによって、その絶対視 (図 1 a 右上) や違和感 (図 1 a 左上) など不適當な誤解を招きかねない (Darling-Hammond, 2013)。今後、アセスメントの目的を形成的評価などの状況改善に置くのであればなおさら、図 1 b のように多様な学問領域や実践現場が諸力を融合して、作問—調査—分析—フィードバックからなる一連の過程を強化することが必要になる。それによって、教科等の本質を問う問題づくりと、その問題解決過程による学習過程の把握、結果を多様な規模・アプローチで分析することを通じた解釈とフィードバックが一貫した形で効果的・効率的に行われることになるためである。

以上のように考えると、この作問—調査—分析—フィードバックというサイクルの強化は、IRT を用いない PBT (Paper-Based Testing) の形式で行われてきた調査においても重要な課題であったと言える。全国学調を例にとれば、それが学校現場の教育指導の充実や学習状況の改善に役立つためには、全国学調の問題が児童生徒の学習の実態に関する課題を問うものとなっており、その結果が実態を明らかにするものとなっており、そこから指導改善に生かせる示唆が得られるものとなっていないとてはならない。しかし、学校現場が「何を問うている問題か」を把握せぬまま、正答率だけを参照し、その向上だけを目指した指導改善に従事するようでは、その目的は実現され難い。すなわち、IRT を用いない PBT においても、上記のサイクルの鍵を同定し、それを強化することが重要だということになる。それがひいては、IRT を用いた CBT (Computer-Based Testing) へのスムーズな移行も可能にすると考えられる。

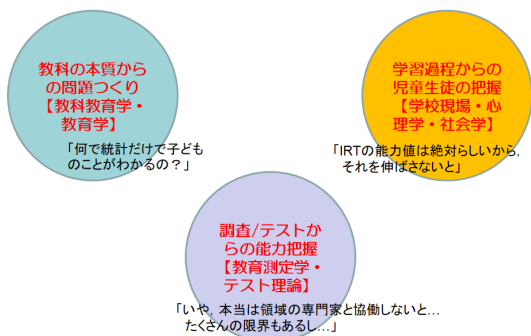


図 1 a. アセスメントを担う関係者間に  
危惧される断絶

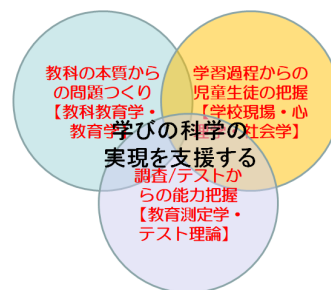


図 1 b. 学習科学の視点からの作問・分析  
サイクル強化による関係者の融合

そのためには、単に学問上の融合を目指すだけでなく、実際に作問に携わる関係者がその業務と密接に連動する形でサイクルを強化できるかに取り組んでみるのが有効だろう。本稿はその課題意識のもとに、全国学調を題材として実施された「学力アセスメントの在り方に関する調査研究：作問・結果分析班」の取組の一部を報告する。

この作問・結果分析班では、学習科学（Learning Sciences）を図1bの諸力融合の中核として活用し、実際に作問・分析に携わる当事者と協働し、ワークショップ形式で研究を進めることで、作問・分析者をエンパワーしながら、研究期間終了後も目指すサイクルが着実に実装されていくことを狙った。ここで学習科学を中核としたのには、二つの理由がある。一つは、学習科学が、認知科学という人がいかに考えるのかなどの認知過程（cognitive process）の解明を目指してきた学問分野から発展してきたためである。二つは、学習科学は解明にとどまらず、常に現場への還元（フィードバック）を目指してきた実践的な分野だからである。

その特徴を生かして、全国学調においても、児童生徒の認知過程の想定と検証を重視することとした。認知過程は、児童生徒が外部の情報に対して、自ら考え理解し判断するといった内的に起きるプロセスである。全国学調の際も、一人一人の児童生徒が認知過程を働かせた結果が解答として表現され、それが解答類型に従って採点されることになる。その作業に当然含まれているはずの「認知過程の事前の想定と事後の検証」を重視することによって、作問・結果分析の担当者にとっては、教科の本質からの作問を児童生徒にいかなる認知過程を引き起こしたいかという観点で行え、結果の分析をいかなる認知過程が引き起こせたかの観点で行うことになる。それが契機となって、作問・実施・分析・フィードバックのサイクル——いわば年度ごとの全国学調における PDCA（Plan-Do-Check-Act）サイクル——を効果的・効率的に回すことにつながるのではないかと作問・結果分析班では考えた。さらに、こうしたサイクルが確立できれば、学校現場にもアセスメントの本来的な目的やその結果の解釈の仕方が伝わりやすくなり、調査関係者自らもサイクルを回しやすくなると期待できる。

以下では、学習科学の考え方に触れながら目指す作問・分析プロセスの概要を述べ、取組の全体像を示す。その上で、取組の核となった「思考発話調査」による調査結果分析の強化に焦点を当て、ワークショップの実例と、実務への反映について報告する。最後に今後の課題を提示する。

## 2. 全国学調の目的達成のためのビジョン

文部科学省は、全国学調の目的を、下記のように整理している（文部科学省，2024）。

- ① 義務教育の機会均等とその水準の維持向上の観点から、全国的な児童生徒の学力や学習状況を把握・分析し、教育施策の成果と課題を検証し、その改善を図る。
- ② 学校における児童生徒への教育指導の充実や学習状況の改善等に役立てる。
- ③ そのような取組を通じて、教育に関する継続的な検証改善サイクルを確立する。

上記①～③の目的は、③の「教育に関する継続的な検証改善サイクル」の確立を見据えて、すなわち、それを全体のゴールとイメージして、PDCA サイクルとしての全国学調企画・立案・実施プロセスの質を向上することで達成される。

その全体ビジョンを例示すれば、図2のようになるだろう。すなわち、国レベルでは学習指導要領の理念に従って、教育と評価の指針が作成され、学校での教育実践を全国学調で検証し、その調

査結果と改善提案を教育現場と共有することで、学習指導要領のより十全な実現を図る（図2 a）。そのサイクルを通じて図の中央にあるような子供の学びとその教育についての見識を高め、国民全体で長期的・持続的に子供たちの資質・能力を伸長することを狙う。

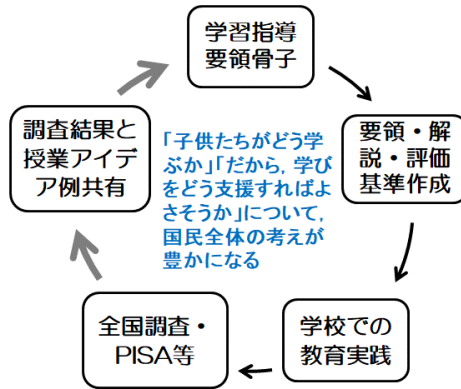


図2 a. 国レベル

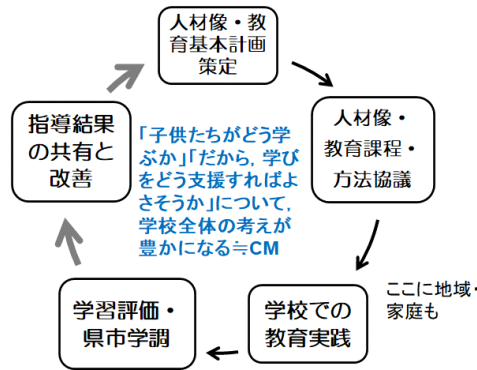


図2 b. 自治体・学校レベル

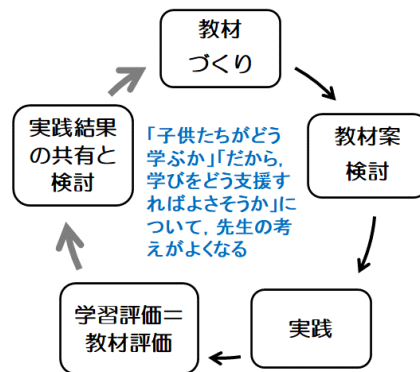


図2 c. 授業（教室）レベル

図2 全国学調の目的達成ビジョン（「CS 調査」は学習指導要領実施状況調査）

これと呼応しながら、自治体・学校レベルでもそれぞれ設定した教育目標に従って、教育課程や方法をデザインし、実践結果をそれぞれの評価方法で検証しながら、結果を共有・吟味して、目標の達成度や次の課題を同定する（図2 b）。このサイクル自体がカリキュラム・マネジメント（以下一部で「CM」と略す）である（文部科学省，2021）。このサイクルを回すことで、自治体・学校における子供の資質・能力育成のビジョン共有とその実現に向けた知見が蓄積されることが全体の狙いとなる。

それぞれの教室でも、各教師が子供にどういう力をつけたいかという目標に従った教材づくりを行い、子供の学びに関する想定を協議して、実践し、教室で学びを見とって、その結果を共有し次の授業改善につなげる(図2c)。それが子供の成長を介した教員自身の成長につながることを狙う。

三つのサイクルはいずれも、各現場の実践と評価の質向上を最大化することを目的としている点、実践と評価の改善は一体的・継続的に行われることを当然の前提として展開されている点で共通している。それらのサイクルにおける評価は自治体や学校、児童生徒の序列化ではなく、それぞれのレベルで行う教育実践が子どもたちの学びにどのように影響したか、学習の実態の把握に主眼が置かれることになる。だとすると、全国学調の結果も、教科ごとの正答率に着目するだけでは、十分でない可能性もある。児童生徒の学習の実態を把握しようとする上では正答/誤答は一観点に過ぎず、正答/誤答に至る思考の過程へのアプローチが鍵となる。例えば、「テストワイズネス」に関する研究でも示されているように、学習者はテストで問題に取り組む際、問題文や解答形式など様々な要素の影響を受けて出題者の想定とは異なる方略を用いて解答を求めようとすることもある(益川ら, 2021)。であれば、国レベルでは、そもそもその問題が何を評価したい問題なのかを今以上に明確にしておくこと、あわせて、本来評価したいポイントで正答、誤答が分かれているのか、それ以外の点につまずきの要素があり、評価したいポイント以外で正答、誤答が分かれてしまっているのかを吟味した上で結果をフィードバックし、次の問題の改善に生かすことが不可欠であると言える。自治体・学校レベルでも、国が示す評価問題の意図と結果を照らし合わせながら、結果の解釈を今以上に深めることが求められる。すなわち、学習状況の把握という観点から国レベルのサイクルを強化した上で、三つのサイクルを同じ観点で一気通貫させることが、全国学調の目的達成のための基本的なビジョンとなる。

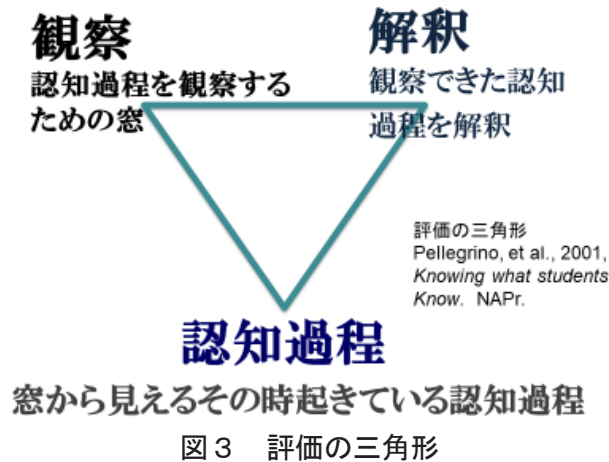
### 3. 学習実態をいかに把握するか—評価の三角形の考え方—

それでは、学習実態の把握に主眼を置いた「評価」とはどのような営みとして整理できるだろうか。一つの基礎となるのが、学習科学の分野で提示された「評価の三角形」(Pellegrino et al., 2001)の考え方である。この考え方は図3に見るように、評価を、把握したい認知過程を事前に想定(モデル化)し、想定に照らした適切な観察を工夫して実施し、その結果を解釈することだと整理している。この考え方は一見当たり前に見える。全国学調で言えば、「観察」の仕方を工夫することは作問、「解釈」は結果の分析に当たると対応付けて納得することもできるだろう。

しかし、そのポイントの一つは「認知過程」を「観察」から切り離れた点にある。学習自体は目に見えない頭や心のはたらき、すなわち「認知過程」であり、直接把握できない。そこで、学習者に問題を出して解いてもらったり、パフォーマンスをさせたり、学習実態の把握の手掛かりとなる知覚可能なデータを得るための「観察」を行う必要があると考えるのである。それゆえ、作問の段階では、児童生徒が問題と出会ったときにどんな資質・能力を発揮して、どんなふう解いていくか、どんなつまずきがありえるかなど、実現しうる認知過程を想定しておくことが重要になる。例えば「長方形の特徴を理解しているとはこういうことだ」といった認知過程のモデルに基づき、学習実態を推し量ることに使える言動データが表出される観察の仕方を工夫するわけである。

二つ目のポイントは「観察」と「解釈」を切り分けた点にある。「この生徒はこういう問題にこのタイミングでこういう解答を書けた/この選択肢を選ばなかった」というデータを得ることは評価の一部でしかない。そのデータをもとに、分析の段階では、実際に想定したとおりの認知過程が生じ

ていたと言ってよさそうか、想定していないプロセスで正答に行きついてしまったり、想定どおりのプロセスで考えても正答に辿りつけずに終わったりといったことがなかったか、想定を一つの基準として認知過程を検証する、すなわち解釈することが重要になってくる。



#### 4. 認知過程の想定と検証を軸とした作問・分析

以上を踏まえ、PDCA サイクルとしての全国学調企画・立案・実施プロセスの望ましい在り方を整理したのが図4である。サイクルは、つながり合う五つの「ステップ」(図中ボックス)からなる。

- 作問のステップでは、各教科等で育成・評価したい力を明確化し
- 問題検討のステップでは、その問いたい力と問題との対応など、問題の詳細について、作問者以外のメンバーが実際に解いてみながら、改善を重ね
- 調査のステップで実施し
- 結果分析のステップで児童生徒の実態を把握することで、各自治体・学校の児童生徒の序列化ではなく、問題そのものの妥当性を検証することを通して
- 調査結果の共有と検討のステップで、検証結果を教育現場に戻して、その教育改善に活用してもらると同時に、翌年度の作問に知見（作問原則など）を活用する。

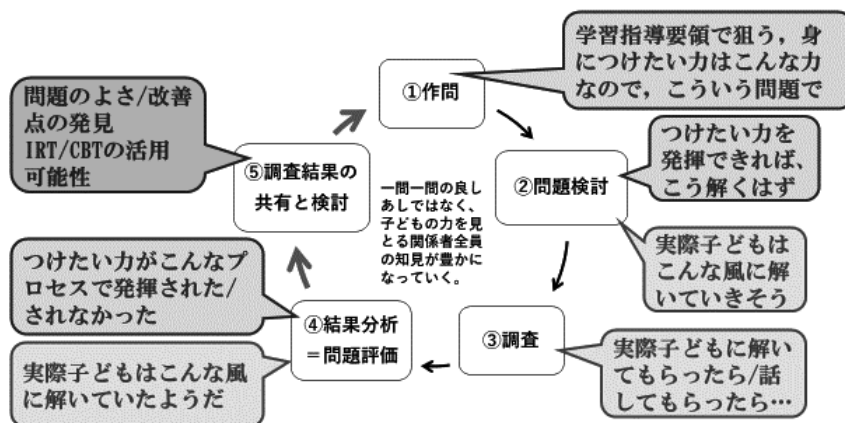


図4 認知過程の想定と検証を軸とした全国学調のPDCA サイクル

## 5. 2年間の研究の展開

図4のサイクルの確立と活用を視野に取組を計画した。各年度に実施したワークショップをPDCAサイクル各段階にあてはめたものが図5と6である。

令和3、4年度の2年間は、令和5年度において、より業務に埋め込まれた形で自律的に回されることを目指し、これまでの蓄積に基づいて実装の見通しを得ることを目的とした。そのため、実施形態も一旦PBTを前提とし、参考情報の位置づけでCBTにも触れる形とした。令和3年度は、実現したいPDCAサイクルのイメージ共有に主眼を置いた。そこで、認知過程に焦点化することの意義が実感されやすいよう、あえて、新規な問題開発（Planの段階）から始めるのではなく、既出問題とその分析結果からスタートし、仮説検証過程も既出問題を活用することにした。令和4年度は、図6に見るように当該年度の全国学調業務を対象とすることによって、実務と連動した「生きたサイクル」としてPDCAサイクルを回していくことを支援するワークショップとした。

### 令和3年度ワークショップ

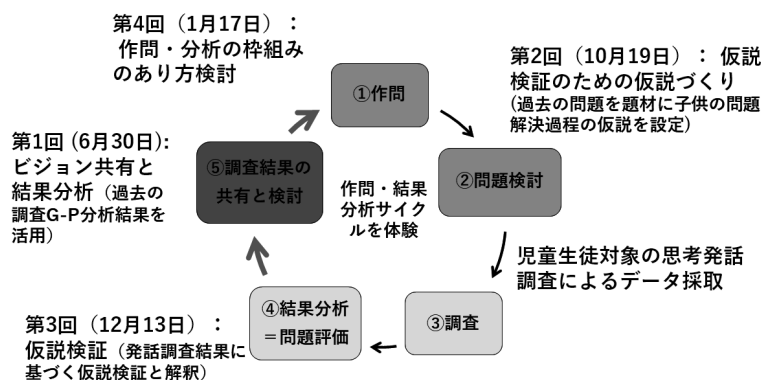


図5 令和3年度ワークショップを当てはめたPDCAサイクル

### 令和4年度実務連動型ワークショップ

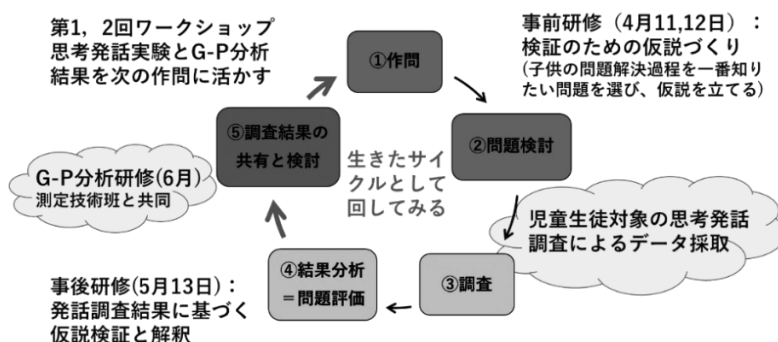


図6 令和4年度ワークショップを当てはめたPDCAサイクル

なお、計画において、サイクルにおける認知過程の想定と検証のための核となる取組として位置づけたのが、「思考発話調査」である。思考発話調査とは、実験参加者に考えていることを口に出しながら問題を解いてもらうことにより、問題解決中の認知過程を推測する手法である。1970-80年代に、問題解決後に自身の認知過程を振り返る活動（内観報告）との対比の観点から、人が問題を

解くときに何を考えているのか、どのように問題を理解するのか、どのような知識を思い出してどう使おうとしているのかといった「高次認知過程」を把握する手法としての有効性や留意点が吟味され、手法が確立された (Ericsson & Simon, 1980)。現在は OECD の PISA 調査において、出題前に、出題者の意図通りに受験者が問題を理解し、意図通りに考えて正解・考えずに不正解にたどり着いているかを確認して作問に生かすといった形で使用されている。そこで、「思考発話調査」の導入は、目指すサイクルの実現の手段となりうると考えた。

ただし、本稿では紙幅の都合で割愛したが、思考発話調査以外に統計データを用いた認知過程の想定と検証作業も別途行った。詳細は報告書を参照されたい。

## 6. 令和3年度ワークショップにおける思考発話調査の意義の実感

### (1) 全国学調過去問題を活用した思考発話調査の試み

令和3年度は、研究者が行った思考発話調査のデータを分析し、問題を評価してみることで、思考発話調査の有益性と認知過程の想定や検証の意義を検討することをワークショップの主題の1つとした。そのために、全国学調の過去問題を簡易 CBT 版にアレンジした問題を題材に令和3年11月15日に公立小学校6年生13名(ソロ5名、ペア4組)対象に調査を実施した。調査者は学習科学研究者など5名が当たり、コロナ禍中であったため、web 会議システムを用いてオンラインで行った。児童は GIGA 端末を活用して調査に参加した。対象とした問題の原案は図7のとおりである。

問題は、「辺が垂直になるように重ねてできた四角形が長方形になっていることを、言葉を使って説明」する問題と、図7中の「図2」「図3」の長方形の辺の長さを一つずつ選ぶ問題とで構成されていた。そのときの調査では、前者の言葉による説明が「全ての角が直角であること」に言及して解答でき、かつ「図2」は選択肢1、「図3」は選択肢2を選択できた児童は全体の9.1%であった。

**9**

次のような、2枚の長方形の紙があります。たろうさんは、この2枚の長方形を重ねてできる四角形について考えます。

9 cm  
3 cm

8 cm  
4 cm

まず、この2枚の長方形を、下の図1のように重ねました。

図1

図1の重ねてできた四角形①が平行四辺形になっていることを、たろうさんは、下のように説明しました。

**たろうさんの説明**

長方形の2組の向かい合う辺は平行なので、重ねてできた四角形①の2組の向かい合う辺も平行になります。

2組の向かい合う辺が平行になっているので、重ねてできた四角形①は平行四辺形になっています。

たろうさんの説明のように、図1の重ねてできた四角形①は平行四辺形になります。

次に、この2枚の長方形を、下の図2や図3のように、辺が垂直になるように重ねたところ、重ねてできた四角形は長方形になりました。

図2

図3

図2や図3のように、辺が垂直になるように重ねてできた四角形が長方形になっていることを、言葉を使って説明しましょう。

また、図2や図3で重ねてできた四角形②や四角形③は、どのような長方形ですか。下の1から4までの中から1つずつ選んで、それぞれ番号を書きましょう。

- 1 縦が3 cm、横が4 cmの長方形
- 2 縦が3 cm、横が8 cmの長方形
- 3 縦が4 cm、横が9 cmの長方形
- 4 縦が8 cm、横が9 cmの長方形

図7 令和3年度小規模思考発話調査問題の原案

<2 ページ目当初画面>

**【問題】** 次のような、2枚の長方形の紙があります。たろうさんは、この2枚の長方形を重ねてできる四角形について考えます。  
まず、この2枚の長方形を下の図1のように重ねました。

(2ページ目)

①

図1

(続きがあります)

<2 ページ目クリック後画面>

**【問題】** 次のような、2枚の長方形の紙があります。たろうさんは、この2枚の長方形を重ねてできる四角形について考えます。  
まず、この2枚の長方形を下の図1のように重ねました。

(2ページ目)

①

図1

(続きがあります)

<3 ページ目画面>

図1の重ねてきた四角形①が平行四辺形になっていることを、たろうさんは、下のように説明しました。

(3ページ目)

**たろうさんの説明**

長方形の2組の向かい合う辺は平行なので、重ねてきた四角形①の2組の向かい合う辺も平行になります。  
2組の向かい合う辺が平行になっているので、重ねてきた四角形①は平行四辺形になっています。

たろうさんの説明のように、図1の重ねてきた四角形①は平行四辺形になっています。

①

図1

(続きがあります)

<4 ページ目当初画面>

次に、この2枚の長方形を、下の図2や図3のように、辺がすいちよくになるように重ねたところ、重ねてきた四角形は長方形になりました。

(4ページ目)

②

図2

③

図3

(続きがあります)

<4 ページ目クリック後画面>

次に、この2枚の長方形を、下の図2や図3のように、辺がすいちよくになるように重ねたところ、重ねてきた四角形は長方形になりました。

(4ページ目)

②

図2

③

図3

(続きがあります)

<5 ページ目画面>

図2や図3のように、辺がすいちよくになるように重ねてきた四角形が長方形になっていることを、言葉を使って説明しましょう。

(5ページ目)

②

図2

③

図3

(続きがあります)

<6 ページ目画面>

図2や図3のように、辺がすいちよくになるように重ねてきた四角形が長方形になっていることを、言葉を使って説明しましょう。

また、図2や図3で重ねてきた四角形②や四角形③は、どのような長方形ですか。下の□にあてはまる長さを考え、解答用紙に書きましょう。

図2の四角形は、たてが□ cm、横が□ cmの長方形です。

図3の四角形は、たてが□ cm、横が□ cmの長方形です。

(問題はこれで終わりです)

②

図2

③

図3

図8 令和3年度小規模思考発話調査問題（Microsoft PowerPoint 画面遷移）

思考発話調査に当たり、第2回ワークショップとしてデータ分析の事前準備を行った。令和3年10月19日に21名が参加し、210分で行った（現況分析班と共同で行ったため、そのうちの160分を使った）。第2回ワークショップでは、算数・数学チームにおいて、この問題での児童生徒の解決プロセスの想定とそれに基づく問題の改善案を協議した。その結果、「もう少しステップバイステップで（受検者に）聞いていけるとよいかも」「四角形を『重ねる』イメージができる」とよい」「問題文の出題の仕方、例えば必ずリード文を読ませる仕組みを工夫したい」などの意見が出された。思

考発話調査ではこれらの意見をふまえ、図8のようにアレンジして出題した。

図8に掲載した「2頁目当初画面」では、クリックすると、赤い長方形と青い長方形が下方にスライドして重なった図形（平行四辺形）を形成するのが見られる。それにより四角形を「重ねる」イメージが動的に見えることになる。その上で3ページ目を提示することで、問題文をいったん読まざるを得ない構成とした。次に4ページ目に再度二つの長方形を提示し、今度は両者が垂直に重なるようにスライド移動した。その上で「言葉で説明する問題」と「どのような長方形か」をステップバイステップで聞いた。加えて、後者は選択肢問題ではなく、長さを直接穴埋め問題の形式で聞くこととした。

これを用いて思考発話調査を行った。本調査の場合は、まずは7分で説明と練習問題を使った発話練習を行い、思考発話そのものに対する慣れと、調査者と調査参加者との間のラポール形成を図った。その後、児童の発話が不十分である場合は、調査者が実演してみせて、求める活動のイメージを共有した。練習後、図8の画面を児童自身でクリックしながら、まずは問題を読み上げ、そこで考えたことを発話していく方法で調査を進めた。時間は最長10分とった。児童が30秒ほど続けて無言でいた場合は、「今何を考えていますか？」と発話を促した。解答については別途紙で配布していた用紙に記入してもらった。また児童はその用紙に自由にメモしてよいこととした。10分経過後、若しくは児童が終了を宣言した後、10分以内で解き方に関する簡単な振り返りをインタビュー形式で行った。

調査対象者がペアの場合も同様に行い、二人の間の対話を推奨し、その対話から一人一人の解決プロセスを把握することとした。発話は全て書き起こし、画面遷移等の行動を付与して、第3回ワークショップ用に整理した。

## （2）思考発話調査に基づく児童生徒の認知過程へのアプローチ

第3回のワークショップは、「問題が引き出す児童の解答プロセスの検証」と題して、令和3年12月13日に120分で21名を集めて行った。目的は、思考発話調査結果に基づく仮説検証とその解釈、仮説の見直しを体験し、認知過程の想定と検証を軸とした作問・分析サイクルについてより深く理解することであった。

プログラムでは、まず狙いを確認した後、思考発話調査を紹介し、実際に児童が調査を受けている動画を、顔にぼかしを入れた状態で共有した。その後、ワークショップ参加者も自分たちで問題を解いてみて内容を確認した。その上で、正答や測りたい力と解答プロセスの想定（図9）を確認した。調査参加者の解答状況を確認した上で、データの説明を受け、発話データの分析を行った。具体的には、グループ内で四つのデータを分担して席を移動して同じデータを分担したメンバーと分析を行い、また元のグループに戻って思考過程の解釈と問題の振り返りを行い、その結果を全体交流した。最後にCBTの本格導入に向けた協議も行ったが、紙幅の関係で詳細は省略する。

上記の「解答状況」は、言葉による説明は、正答が1名、誤答が4名、無答が8名であった。また辺の長さは、正答が9名、誤答が3名、無答が1名であった。提供した発話データは、図10のとおりであった。左がソロ、右がペアでの調査である。データは、左列から行数、開始からの経過時間、活動のおおよそのフェイズ、閲覧画面ページ、調査参加者（図中「実験参加者」）の言動メモ、調査者（図中「実験者」）の働きかけを示している。これを見ると、どういう順番で児童が問題を読み、どんな情報を使って解いていったか、どんなことを考えていたかを一覧できる。例えば、児童が4ページと6ページを見比べながら長さを見つける問題を先に解こうとしていること、解決に必

要な横の辺の長さは2ページにしか掲載されていないにもかかわらず(図8参照)、なかなかそのページに戻らないことが見てとれる(右側のペアは19行目に2ページに戻って初めて「あった」と発言している)。

下記が分析結果の全体交流で出された意見である(原文ママ。カッコは引用者)。いずれも「子供の問題の解き方に関する気づき」が、矢印以降の「問題の作り方に向けての気づき」とつながる構造となっている。

【資料4】思考発話調査における解答プロセスの想定		題材：2枚の長方形の紙を重ねてできる四角形問題 (小学校算数過去問を参考にした簡易型 CBT)	
1. この調査問題で評価したい力 (過去問題の解説に基づいて)			
直線の平行や垂直の関係を基に、重なってできた図形と長方形の性質を関連づけることができる			
2. 児童生徒の解答プロセスの想定			
下で想定した解答プロセスのうち、実際に多く観察されそうなのはどれでしょう。また、他にありそうなものとしてはどんなプロセスが考えられるでしょうか。			
参考【もとの問題の反応率】			
・ 説明&数値とも正解…8.6%			
・ 「すべての角が直角」という長方形の定義以外の特徴に触れた説明を書いたが、数値は正解…28.4%			
・ 長方形の特徴をふまえない説明および説明は無回答で数値は正解…33.2%			
○正答を選ぶ場合の典型的な過程 (思考の具体例)			
1	問題理解にあたって - 問題を頭から読み、冒頭の説明文と図、図1の動(様子)から、「[3cm×9cm]の長方形の紙と[4cm×8cm]の長方形の紙を重ねて四角形をつくる」という前提状況を読み取る。 - 図1と「たろうさんの説明」から、四角形を重ねて平行四辺形ができた場合を例に、「2枚の紙を重ねてきた四角形が特別な四角形になっていることについて言葉を使って説明する」とはどういうことかのイメージをつかむ。 - 2ページ目の説明文と図2、図3の動(様子)から「[2cm×9cm]の長方形と[4cm×8cm]の長方形を辺が垂直になるように重ねると、(互い違いに重ねても、同じ向きで重ねても)重ねてきた四角形は長方形になる」という問題状況を読み取る。 - この問題状況について、「図2や図3のように紙を重ねてきた四角形(②・③)が長方形になっていることを、言葉で説明すること」および「②や③の四角形の縦と横の長さを同定すること」の2つを求められていることを理解する。	解法の探索・実行にあたって - 「長方形」の定義(すべての角が直角)を想起する。 - 4・5ページを見直しながら、2枚の長方形の紙を辺が垂直になるように重ねると、重なった部分の角がすべて直角になっていることを確認する。 - 3ページの「たろうさんの説明」の書き方を参考に(～なので、重ねてきた四角形②や③は…になっています)、言葉を使った説明を考える。 →「辺が垂直になるように重ねてきた四角形はすべての角が直角になっているので、四角形②や③は長方形になっています」と解答用紙に記載する。 - 図2を見て、②の縦の長さ(縦)と赤い紙の短辺の長さが同じであり、横の長さ(横)と青い紙の短辺の長さが同じであることに気づき、数値を探しに戻る。 →4ページを見て、四角形②は縦2cm横4cmであると判断する。 - 図3を見て、③の縦の長さ(縦)と赤い紙の短辺の長さが同じであり、横の長さは青い紙の長辺の長さと同じであることに気づく。 →2ページを見て、四角形③は縦2cm横8cmであると判断する。	行き着いた解答 「辺が垂直になるように重ねてきた四角形はすべての角が直角になっているので、四角形②や③は長方形になっています。」  四角形②→ たて2cm横4cm 四角形③ →たて2cm横8cm

図9 児童の解決過程想定のためのワークシート(一部)

時間	活動	閲覧ページ					実験参加者 活動	実験者
		P.2	P.3	P.4	P.5	P.6		
1	0:12:19	開					問題開始	
2	0:12:45	開					読み上げ	
3	0:13:22	開		1				
4	0:13:55	開				1		
5	0:14:10	開					1	
6	0:14:40	開					読み上げ終わり、問題解決し始める	
7	0:16:06	開			1		「図2って赤い方の縦と横？」	「説明がないかな？」
8	0:16:48	開				1	「縦はいいですか？」	「どう書いてあったかな？」
9	0:16:56	開				1	「図2の青と赤が重なっているところ？」	
10	0:17:07	開				1	ページを探す	
11	0:18:55	開				1	「これか! 図1の説明をする「平行四辺形のところとか、青と赤の重なった部分を探る感じか」	「今、何理解した？」
12	0:18:59	開				1	「長方形になっていることをどこで説明しましょう」	
13	0:19:13	開				1	□cmの数字を入れながら長方形の説明を色々する。図3の横がわからず、考えている。	
14	0:24:21	開				1	分からない横の長さをどうにかして考えている	「あと3分です」
15	0:25:46	開				1		「あと1分です」
16	0:27:01	開				1	「どっちも長方形で、青も赤も長方形だから、多分そこが垂直、全部垂直になっているから、重なったところも同じ感じになると思うから、多分長方形になるのかな」	「6ページ目これはなんて説明する？」
17	0:27:19	開				1	解答用紙の提出	「ありがとうございました」
18	0:28:54	開				1	「この赤と青の分を図3のようにしたら、図3のように重ねたら、交わる場所というか、紫の部分か、二つ、多分、まっちゃんできる？」	インタビュー「図3の横の長さについてどう考えた？」
19	0:29:06	開				1		
20	0:29:26	開				1		

図10 思考発話調査のデータ例(左ソロ、右ペア)

問題のよしあしを直接的に議論するのではなく、児童の問題の解き方(認知過程)を踏まえて、不要なつまづきの可能性を減らすなど、児童の力を引き出した上で、最も問いたい知識等の有無や

思考力等の発揮度合いを見極めるにはどうしたらよいかという形で問題の作り方を考える構造だと言える。

- そもそも重なっている部分がどこかが分からなくてつまずいたり、2ページで書かれている横の8cmの情報が見つけれずにつまずいている子が多かった。「状況に入れば」が大きなポイント。  
⇒状況に入るための情報はしっかりと明確に提供していく方が問いたい能力が問える（「状況に入る」とは、児童・生徒が問題状況を理解し、問題の提示する文脈に沿って思考し始めることを意味していると解釈できる）。
- 閲覧ページを見ると問題が出ている後ろの方を集中的に見ていて、問題全体を見ているのが少ない。  
⇒なので、できるだけ問題はコンパクトに収めていくと良い。子供は「何を聞かれているか」に集中して考える。  
⇒今作問するとしたら、たろうさんの説明の後にすぐ長方形の説明を求める。「たろうさんの説明と同じように」など、どういう説明をしたらいいかが子供に分かるように。長さの話も分けて聞く。
- 子供が「長方形について説明してください」と言ったときに「縦と横の長さが違うから」といった説明をするのは、2年生のときに正方形と長方形と一緒に習っているので。実験者（注：調査者）が「その説明だったら平行四辺形もそうじゃん」って言って、初めて角度に着目してくれている。  
⇒子供が「なぜ長方形の説明をするときに直角という言葉が必要か」が分からないときに、そのことを引き出す工夫が要る。

以上、図4と対応づければ、問題の調査(D)を思考発話法で行い、その結果を分析して(C)、次のアクションへとつなげていくワークショップとなった。加えて、調査のために特別な問題を作り込まなくても、既存の全国学調の問題の解決過程で児童から直接得られるデータは、その認知過程の特徴とそれを踏まえた作問の工夫について関係者の考えを深める効果があることも示唆された。

## 7. 令和4年度ワークショップにおける思考発話調査と実務の連動

令和4年度は、令和3年度の実践を踏まえ、全国学調において作問と結果分析を中心としたPDCAサイクルを回しつつ、それを全国学調の当該年度の問題に関する報告書作成など、実務と連動させる形で実装していくことに取り組んだ。実装の対象になったのが、思考発話調査である。令和4年度の調査は、事前研修、協力校での調査（＝データ採取）、データの分析視点を示し調査結果をいかに報告書に活用しうるかなどを検討した事後研修、調査データ分析ワークショップのセットで実施され、結果は関係者の判断により適宜、全国学調の報告書や夏の全国説明会に反映された。

事前研修は、令和4年4月11日及び12日の二回に分けて、調査実施担当者計60名を対象に行った。研修内容は、思考発話調査の狙いや調査方法、留意点の説明及び議論と、各教科に分かれての対象問題の確定、調査方法の確認である。調査方法と留意点の確認のためには、するべき発話ややるべき作業を明確化した「手順書」（図11）と、調査で見とりたいポイントを明確にし、期待するプロセス／期待しないプロセスと正答／誤答の関係を整理する「見とりの観点シート」（図12）を各教科チームで作成・共有することとした。こうした書式は前年度のワークショップとも共通するものとなっており、過去問題を使ったワークショップから、現在進行形の実務と連動したワークショップ、そして実際の実務へと、「認知過程の想定と検証」の視点が着実に受け継がれ、実装されてい

く媒介として機能することを狙って作成した。

(1) ソロの場合

時間	実験者から児童への教示内容等	児童の活動	配布物と支援
1分 教示 ( : ~ : )	「では、始めましょう」 「やっている途中で思い浮かんだことは、できるだけどどん口に出してください」 「こちらからも質問をすることがあります。「なんとなく」でもいいので、できる範囲で教えてください」 「解答は解答用紙に記入してください。メモ欄も自由につかってください」 「時間はX分です。3分前になったらお知らせします」	説明を聞く	本番問題解答用紙 配布した状態から
X分 問題解答・思考発話 ※無言&個人で解いた場合の倍程度の時間を想定 ( : ~ : )	①では、スタート。まずは問題を声に出して読んでみましょう。  ②解いているとき（無言の場合、適宜以下の声掛け） 「今何を考えていますか？」  ③3分前 「あと3分です」	思考発話で問題を解く	ビデオ、ICRの記録支援

図 11 思考発話調査の手順書ひな形（一部）

思考発話調査のための見とりの観点シート（付：記入例「あた」）

今回の調査で見とりたいポイント	この問題で見とりたい（詳細したい）学びのプロセス（認知過程）	期待と通りのプロセスで正答に行き着く姿の具体例（発言など）	期待と通りのプロセスではないかもしれないが、正答してしまう姿の具体例（発言など）	期待と通りに近いプロセスが起きているものの、正答に行き着かない姿の具体例（発言など）
「情報を解釈する力」（問題のねらい）を発揮できているか	一あなたは指を伸ばした長さ 一あなたは一あなたの1.5倍 という2つの情報を問題文と表から読み取って、図と対応する	〇あなたは指を伸ばした長さで、その1.5倍なので、あと半分のところまでの図だから…4かな	〇数字とか入ってない図が答えっぽいから1か4で…1はなんとなく違うかも	〇1つと半分、指1つと半分だけ…、「一あなた」って書いてなくていいのかな？（答えは空欄）
「言葉や式を使って書きましょう」の指示をどう受け取っているか	言葉と式を使って、答えの根拠を表現できる	一あなたは身長10%だから、妹の場合、 $140 \times 0.1 = 14$ で、妹の「一あなた」は14cm。使いやすいはしの長さは一あなた半だから、 $14 \times 1.5 = 21$ で一あなた半は21cmになる。	〇言葉や式で、倍は掛け算。10%と1.5倍ができたから、かけ算書けばいいよね（ $140 \times 0.1 \times 1.5 = 21$ ）と解答	〇まず身長の10%は、 $140 \times 10 / 100 = 14$ で、その1.5倍だけだ、え、言葉でどういうこと？言葉…（答えだけ記入）

図 12 思考発話調査用の見とりの観点シート（一部）

思考発話調査は令和4年4月22日に公立小学校、4月26日に公立中学校各1校で実施した。調査は小学校25名、中学校38名を対象に、それぞれ45分の1コマを用いて行った。事後研修は、令和4年5月13日午前中学校調査関係者17名、午後小学校調査関係者15名対象に各90分で行った。事後研修では、「協議1：調査者の当日の気づきの共有から、問題を解く過程での子供の思考や理解について見えてきたこと」及び、「協議2：協議1で見えてきた児童生徒の思考や理解の様子を根拠に、報告書を強化したいとしたら、発話をどう分析するか」という二つのテーマで協議を行った。協議1では、思考発話調査を通しての気づきとして、下記のような点が共有された。

- よきにつけ悪しきにつけ既有知識の影響の強さ
- 読むことの難しさ、「冒頭から順に全部読むだろう」との出題者側の想定とのずれ
- 問題解釈の難しさ、「問題と正対すること」に関する想定とのずれ
- 小学校における表面的・形式的・状況依存的解法による正答、それが中学校だと通用しなくなる難しさやギャップ

さらに、令和4年6月28日に小中学校の分析結果を17名で持ち寄り、共有吟味するワークショップを120分間掛けて行った。例えば、小学校国語のチームでは、分析結果から見てきた問題解決のプロセスを図13のようなフローに整理していた。

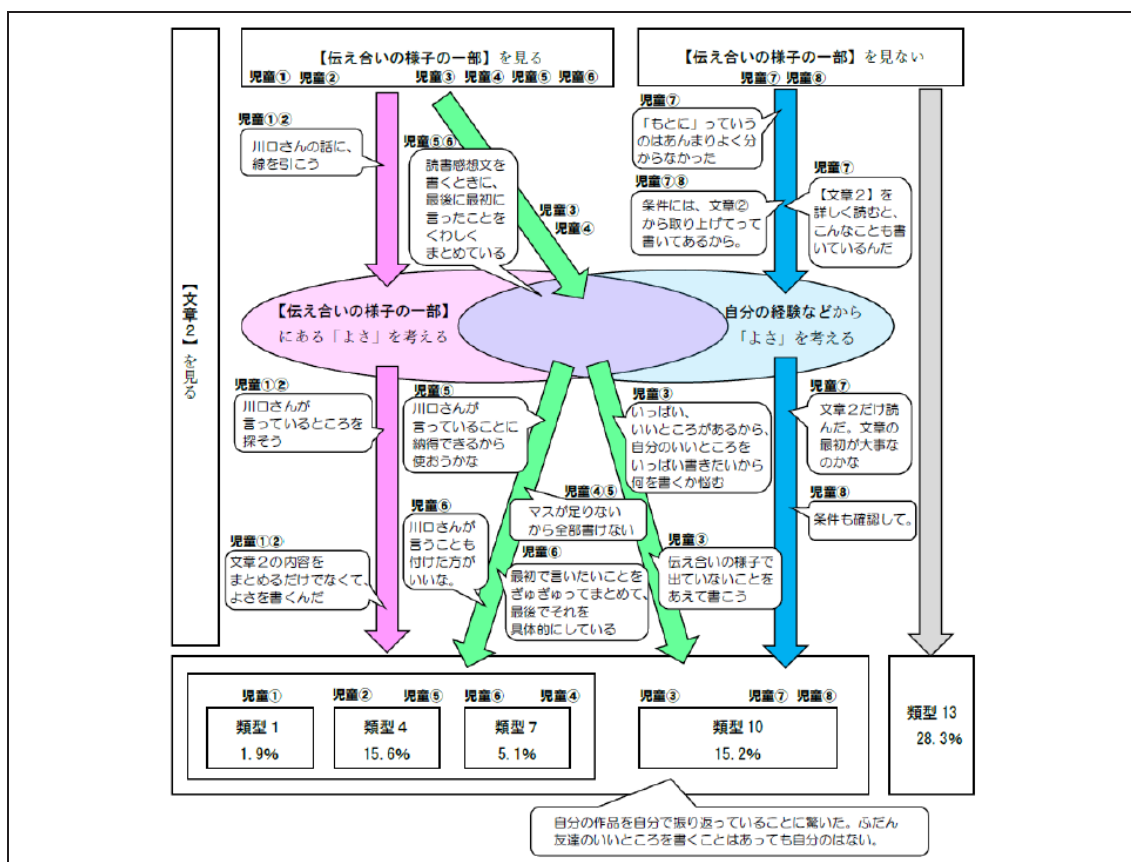


図 13 思考発話調査結果における児童の問題解決フローの整理図（小学校国語）

小学校国語で調査対象とした問題は、「文章2」のよさを振り返って書くというものだが、その際に「伝え合ったことをもとに」という条件が付与されていた。図を見ると、全員「文章2」は読んでいるものの、児童によっては「伝え合いの様子の一部」を見ている児童とそうでない児童に分かれていることがわかる。それでも、正答である解答タイプ1, 4, 7, 10にたどり着いている。その背景に、「伝え合いの様子の一部」から「よさ」を抽出しなくとも、「自分の経験など」から「よさ」を考えるというプロセスがあった。また、「伝え合いの様子」を読んだ場合も、それだけから考える児童と、「自分の経験」を更に加味する児童がいた。こうした多様なプロセスは、思考発話調査から初めて見えてきたものであるという。

国語科では、このフロー図を、正答率や解答タイプごとの反応率などを踏まえて、より抽象化してまとめたものを全国説明会資料に活用した。思考発話調査をとおして、「与えられた情報（この場合は文章）」を基にして「あなたならどう考えるか」を問うタイプの（やや出題趣旨が複雑な）問題における児童生徒の思考過程のパターンを提示することで、どのような考え方をしてほしいかの再考を促し、今後の指導改善の検討に資する資料提供が可能になった例と言えるだろう。

## 8. 成果と課題

以上、「思考発話調査」による作問・分析サイクルの強化を中心に、「学力アセスメントの在り方に関する調査研究：作問・結果分析班」について報告した。

最後に、成果と課題を整理する。児童生徒の認知過程の想定と検証を軸とした作問・結果分析のPDCAサイクルの回し方と意義が、関係者に理解・実践されるようになったことは、本研究のもっとも大きな成果である。7節に示した思考発話調査による全国説明会資料の例（図13参照）にこの成果は端的に表れていると考えられる。作問や結果分析の最も基本的な枠組みについては、既に全国学調の「解説資料」や「報告書」に、「調査問題一覧表」として、「問題の概要」「出題の趣旨」「学習指導要領の領域」「評価の観点」「問題形式」が一覧しやすいように、教科共通で掲載されている。本取組では、思考発話調査を一つの手段として、この枠組みに記載されている項目が実際にどのように問題に具体化され、どのように児童生徒の解決過程に反映され、結果が解釈可能になるのかを関係者の間で十分検討することができた。この成果は、IRTを利用したCBTの開発において、類似問題の大量作成やそのうちの一部のサンプル提供による教育現場の指導改善等の新たな課題に取り組む基盤となりうる。

例えば、出題者が児童生徒の認知過程の観点からどのような「問題群」がある資質・能力を問うものになり得るかを把握できていれば、それだけ作問が効率的に行える。同様に教育現場が同上の観点で結果を解釈できるのであれば、たとえ公開問題が一部であっても被覆された問題を推察できることになる。こうした改善は、テスト理論・教育測定学者に識別力などの従来の指標やアプローチを超えた「問題特性の把握」を求めることになり、今後の課題に取り組む基盤を形成していくことにつながる。

その一方で、2節で示したように、国レベルの作問・分析プロセスの改善は、自治体・学校・現場の教師による指導と評価の質向上と融合することで初めて実質的に機能する。教師や教育行政関係者一人一人が、全国学調を日常の指導評価の一環と位置付け、自分たちなりの問題検討、結果の解釈を踏まえて、子どもの学び方の実態を見とり、授業デザインのアイデアを得るという基盤の充実が不十分であると、IRTのような新しい要素技術が導入されても、それが有効に機能しない可能性もある。

対して、両者が融合的に機能すると、調査結果を基に教師が「なぜ子供がそう解いたのか、間違ったのか」という理由（WHY）を洞察でき、それを基に「だったら授業をこう変えよう」という手（HOW）を見つけやすくなる。この気づきを問題プールに吸い上げていくことができれば、現場の教師が授業づくりと評価の両方に関わる継続的な授業・作問改善プロセスを教育現場全体で創り上げていくことも可能となる。本研究の成果を、国から学校現場まで一貫通貫した指導と評価の質向上の取組に結びつけていくことが今後の課題となる。

### 参考文献

- Darling-Hammond, L. (2013). *Getting teacher evaluation right: What really matters for effectiveness and improvement*. Teacher College Press. (リンダ・ダーリング・ハモンド著；無藤隆・松井愛奈訳 (2024). 『教師に正しい評価を』. 新曜社)
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87 (3), 215–251.
- 益川弘如・白水始・齊藤萌木・飯窪真也・天野拓也 (2021). 「積極的読み」を引き出すCBT読解問題の開発—東京大学入学試験の国語問題を活用して—. *日本テスト学会誌*, 17 (1), 25-44.

文部科学省 (2021) カリキュラム・マネジメントについて. 文部科学省 HP.

[https://www.mext.go.jp/a\\_menu/shotou/new-cs/1384661.htm](https://www.mext.go.jp/a_menu/shotou/new-cs/1384661.htm) (2025 年 2 月閲覧)

文部科学省 (2024) 令和 7 年度全国学力・学習状況調査リーフレット. 文部科学省 HP.

[https://www.mext.go.jp/a\\_menu/shotou/gakuryoku-chousa/zenkoku/mext\\_02942.html](https://www.mext.go.jp/a_menu/shotou/gakuryoku-chousa/zenkoku/mext_02942.html) (2025 年 2 月閲覧)

Pellegrino, J.W., Chudowsky, N., & Glaser, R. (2001). *Knowing what students Know: the science and design of educational assessment*. Washington, DC: National Academies Press.

Scardamalia, M., Bransford, J., Kozma, R., & Quellmalz, E. (2012). New assessments and environments for knowledge building.

In P. Griffin, B. McGaw, & E. Care (Eds.), *Assessment and Teaching of 21st Century Skills*. NY: Springer-Verlag, 231-300.

(スカルダマリアら (2014). 「知識構築のための新たな評価と学習環境」. 三宅なほみ監訳 益川弘如・望月俊男  
訳 (2014). 『21 世紀型スキル-学びと評価の新たなかたち』. 京都: 北大路書房, 77-157.)