

# 学研教育総研レポート

学研教育総合研究所 田村 尚志

## 全国学力テスト 昨年・今年・そして未来へ

4月18日、小学校6年生と中学校3年生を対象に令和6年度の全国学力テストが行われた。今回は、まず、昨年2023年度の結果を一部深掘りし、各都道府県順位だけでは見えてこなかった正答率上位の都道府県の特徴を探ってみよう。そのあと、今年度の問題や実施形態などのトピックスを整理し、さらに、来年度以降実施が予定されている、今とは異なる新しい理論で実施されるテストについて、特徴などを概観してみる。

### ◆昨年度のトピックス

昨年度の一番大きなトピックスは、中学の英語か。まず、その実施形態。「話す」問題がタブレット端末などを使い、オンラインで行われたこと。一部でのトラブルも報道されたが、さらに「話す」結果も話題になった。「読む・書く」の正答率が46・1%と50%を切ったが、「話す」はさらに低く、12・4%で、ある問題は、正答率4・2%だったこと、そして都道府県での正答率の差が大きかったことなど、英語に関しての報道が目立った。また、正答率の高かった自治体で、対策授

業や、過去問を使った事前対策など「やりすぎ」が問われたこともニュースになった。

毎年、正答率の順位で上位にくるある都道府県でも、行き過ぎた事前対策が行われていると指摘された。都道府県教育委員会が実態調査を行い、各市町の教育委員会に向けて、「4月の新学期初日から18日火曜日までの期間は過去問題や類似問題を授業や補充学習で取り扱わないこと」という通知を出したことが報道された。

本来の趣旨と異なるとはいえ、各都道府県ごとに正答率が公開されれば、少しでも点数を上げたくなるのが人情か。

事前対策で報道で取りあげられた都道府県を含む、平均正答率の高かった上位の都道府県の結果を少し深掘りしてみよう。

### ◆昨年度の結果を深掘り

まず、小学校・国語から(図1)。

小学校・国語の正答率上位は、1位秋田、同率1位石川、3位福井、4位京都、同率4位青森。この5府県の正答率を、

①学習指導要領の内容の、思考力、判断力、

図1 小学校・国語

		1位	1位	3位	4位	4位	(%)
		秋田	石川	福井	京都	青森	全国
思考力、判断力、表現力等	A 話すこと・聞くこと	80.1	79.3	77.7	76.2	77.6	72.6
	B 書くこと	30.1	35.7	32.7	29.9	29.0	26.7
	C 読むこと	76.6	75.4	74.6	73.2	73.4	71.2
問題形式	選択式	77.2	76.9	76.9	75.9	74.8	73.6
	短答式	67.6	66.7	65.5	65.1	68.6	62.7
	記述式	58.2	59.7	56.8	54.9	56.3	51.1
記述式の無回答率		13.7	19.4	17.0	20.4	15.5	29.9

国立教育政策研究所・令和5年度全国学力・学習状況調査をもとに教育総研で作成

表現力等の「A話すこと・聞くこと」「B書くこと」「C読むこと」  
 ②問題形式の「選択式」「短答式」「記述式」  
 ③記述式の無回答率  
 のそれぞれで、全国平均と比較してみよう。  
 ①石川の「B書くこと」の全国平均との差が9・0ポイントと、ダントツに大きいのが目

図2 小学校・算数

		1位	1位	3位	4位	(%)
		東京	石川	福井	秋田	全国
問題形式	選択式	63.2	63.4	60.3	60.7	57.7
	短答式	78.0	78.2	78.0	76.3	74.7
	記述式	52.4	51.2	52.1	50.9	47.3
記述式の無回答率		28.4	17.5	16.2	15.4	25.2

国立教育政策研究所・令和5年度全国学力・学習状況調査をもとに教育総研で作成。秋田のほか富山、京都、高知も同率4位

立つ。  
 ②問題形式の「記述式」の差が8・6ポイント大きいのも特徴的。記述式については、京都をのぞく4県が5ポイント以上差をつけているのにも注目。  
 ③記述式の無回答率は、秋田が16・2ポイント差、青森が14・4ポイント差と、頭ひとつ抜けているが、ほかにも無回答率の低さがうかがえる。  
 次に、小学校・算数(図2)。  
 小学校・算数の上位は、1位東京、同率1位石川、3位福井、4位秋田ほか1府2県。

①問題形式「選択式」「短答式」「記述式」  
 ②記述式の無回答率  
 の2つを全国平均と比較すると、  
 ①東京が全項目で全国平均を大きく上回っているが、特に「記述式」で5・1ポイント上回っている。  
 ②記述式の無回答率では、石川、福井、秋田がいずれも10ポイント近い差をつけた低い数字なのに対して、東京は逆に、全国平均より高い無回答率となっている。  
 記述式の平均点が高いのに、無回答率が高いということは、高得点層と低得点層に二極化していることが考えられる。  
 毎年上位にランクされる石川、福井、秋田は、記述式の得点が高いこと、記述での無回答率が低いことなどが他県との違いで、高い得点率につながっているようだ。これは、直前に過去問を解いただけではない、ある程度時間をかけて培ってきたものと見ることができらる。  
 これらの上位常連が今年ほどのような結果になるのだろうか。  
**◆今年度のトピックス①**  
**「情報リテラシー」**  
 小学校・中学校とも、国語の出題内容に特徴があった。  
 中学校・国語では、「フィルターバブル現象」について話し合う場面が設定された。  
 「フィルターバブル現象」について、イン

ターネットで検索・閲覧する履歴から、自分の好む情報だけに囲まれる、あたかもバブルの中にいるような状態を図を交えて説明し、インターネットだけではなく、日常生活の中でも起こり得るという話に展開。本を選ぶときも、自分の好みだけで選ぶと、同じような偏りが生じる……という話し合いをもとに、自分の意見を書かせたりする内容。  
 小学校・国語では、テーマこそ学校での取組となつているが、異なる地域にある2つの学校をオンラインで結び、話し合いを行うという設定だった。  
 いずれも、「情報」を意識した問題で、IGAの中で、ICTの具体的な活用が前提。先に述べる、学力テストが紙をベースとしたPBT方式(Paper Based Testing)からコンピュータベースのCBT方式(Computer Based Testing)へ移行していくことを踏まえたものとも感じられる。  
**◆今年度のトピックス②**  
**「実生活の課題解決」**  
 中学校・数学では、プログラミングカーの停止位置がプログラムした数値から少しずれることから、実際に試した停止位置のデータを集計し、箱ひげ図で比較するという、不確定な事象についてデータをもとに考察する問題や、ストーブの使用量と灯油の残量という日常の題材から関数関係を見だし、数学的に処理する問題などが出題。

小学校・算数でも、桜の開花予想について、過去のデータを分析し、表やグラフ、式に表し、考察する問題が出題された。

いずれも、昨年OECDが発表したPISA Aで日本の課題とされた「実生活における課題を数学で解決する自信が低い」「数学的思考育成のため、日常生活とからめた指導を行っている傾向が低い」の現状が確認できる問題ということかもしれない。

### ◆今年度のトピックス③ 「学習状況調査オンライン実施」

児童・生徒の学習状況調査は、フルオンラインで実施となった。学校ごとの状況に応じて4月10日から4月30日までの期間で実施できた。

そして、教科調査が行われなかった理科・英語の質問も入る内容になっていた。では、次に来年度以降の学力テストがどのようなに変わっているかを見てみよう。

### ◆中学校・理科はCBTで実施

2025年度には、中学校・理科でCBT方式が採用される予定。

CBT方式になることで、紙という制約がなくなり、映像・音声を利用したり、表計算機能などを使う出題が可能になったりする。幅広い内容の出題や多面的な測定ができるようになる。

実験・観察に関する内容は、その手順、結

果とも、変化の様子がきわめて重要だが、映像で示される現象や状況をそのまま提示して、観点を指摘するような出題も可能と、より、授業や学習内容に即した形での実施が期待される。

### ◆古典的テスト理論から 項目反応理論に基づくテストに

CBT方式への移行で、問題形式や内容の変更だけでなく、結果の解釈の仕方が変わる。今までのテストは「古典的テスト理論」に基づくテストだったが、「項目反応理論（IRT (Item Response Theory)）」に基づくテストとなる。

「古典的テスト理論」によるテストは、入試問題をはじめ、今でも多くの場面で使われている正答率や総得点による評価。受験者の能力を集団内での相対的な尺度でしか評価できないが、入試はある決まった集団の中で評価できればよいので、この形でも機能する。ただ、当て推量や運などの要素は排除できないし、問題の難易度をコントロールすることが難しく、異なるテストによる得点を比較することができない。したがって、経年で行われる全国学力テストにおいても、実施された年度内での評価しかできず、例えば、ある教科の点数が前年より上がったとしても、学力が伸びたのか、問題がやさしくなったのかを評価することはできない。

一方、「項目反応理論」によるテストは、

受験者の能力と試験の問題を切り離して、それぞれ別々に評価できるため、問題の難易度を評価した上で、学力が伸びたかどうかを評価することが可能だ。問題の難易度を評価していれば、同じ問題でなくても、受験者の能力を評価できるため、異なる問題を用意し、別々の日に受験することも可能になる。

ただし、いくつかのパラメーターを設定し、確率論から得点を推定するため、多くのデータが必要となる。

また、原則問題は非公開とすることが必要など、制約もあるが、現在、TOEFLやITパスポート試験などは、この項目反応理論に基づくテストになっている。

### ◆競うテストから本来の測るテストへ

「一人一人の学力・学習状況に応じた学習指導の改善・充実」のための学力テストのはずだが、冒頭にみたように競うような側面も払拭されていない状況が、本来の目的になる大きな転機になると期待をもって見守りたい。

