

## Chapter 14 測定している？：多肢選択式問題

(MEASURING UP?: Multiple-Choice Questions)

Ercikan, K. & Seixas, P. (2015). *New Directions In Assessing Historical Thinking*, pp.221-232.

### 著者情報

★ガブリエル・ライヒ (Gabriel A. Reich)

**所属**：バージニア・コモンウェルス大学 (Virginia Commonwealth University)  
中等社会科・歴史教育の准教授

**研究**：評価 (assessment), カリキュラム, 集合的記憶 (collective memory)

**論文**：Testing historical knowledge: Standards, multiple-choice questions and student reasoning,  
*Theory and Research in Social Education*, 37(3), 2009, pp.325-360.

(大学 HP 参照：<https://soe.vcu.edu/directory/full-directory/first--last-name-302962-en.html>)



### 本章で登場する専門用語

- Multiple-Choice Questions：多肢選択式問題 (本章では、多肢から1つの正解を選択)
- think aloud：思考発話法 (問題に解答する際の思考過程を発話させる研究手法)
- dispositional orientation：知的性向 (本章では、歴史の言語に向かう姿勢を指す)
- Test-Wiseness：テスト習れ

### 本章に関連して深めたい論点

①【assessment of learning の観点から】進学や就職のために使用されるハイスティクスな評価は、受験者である児童・生徒はもちろん、その背後に居る家族や地域社会の市民の目にも触れる。そこで、配慮しなければならない歴史のコンテンツには、例えば何があげられるか？それらは、多肢選択式問題として問えるか？ (関連：p.221・p.229)

②【assessment for learning の観点から】多肢選択式問題それだけでは、推論の証拠を得ることができない。かと言って、本章のように思考発話法やインタビューを毎回できるわけでもない (何十人という規模になれば、なおさら)。教育現場で経常的にできる形成的アセスメントとしての多肢選択式問題には、どのような方法が考えられるか？

③【assessment as learning の観点から】テストリテラシー (評価的知識/思考) の育成を加味した歴史試験問題をどのように作成することができるか？そもそも、教師・生徒は「テスト」という機会を学習の場と見なしているか？ (関連：p.228)

## 本章の概要

### 1. 歴史の達成モデル (Models of Achievement in History)

●歴史教育研究界では、実証研究に基づいた歴史のディシプリンに関する理解の達成モデルが複数存在し、学者コミュニティの中での緩いコンセンサスがある (Lee, 2005; Lévesque, 2008; Seixas, 1996; Wineburg, 2001)。

●他方、各国・地域による公式な内容スタンダード (official content standards) も歴史のディシプリンに関する理解の達成モデルの一つとして見ることができる。公式な内容スタンダードは、歴史教育界の影響を受けているかもしれないが、制度的文脈の中で開発されるものであり、義務や政治的な考慮事項に左右される。加えて、市民の目にも晒されている。→教育官僚が歴史の内容スタンダードを定義することは、政治的プロセスとして遺産や文化の次世代への継承に対する世間の不安・期待に対処することを意味し、異議があれば国民に説明責任を果たさなければならない。

●本章では、1999年にニューヨーク州教育局 (NYSED) が開発した「コア・カリキュラム」と「スタンダードとパフォーマンス指標」を反映した多肢選択式テストに焦点を当てている。

★「コア・カリキュラム」…高校2年までに教師がカバーすべき歴史のコンテンツのリスト

→「マーシャルプラン」のような事実に基づく知識と「超大国の出現」のような物語的知識から構成

★「スタンダードとパフォーマンス指標」…歴史教育で育成すべき概念的理解と歴史的思考の基準

●受験者が一連の学習基準を満たしているかどうかを観察するためには、実際に学習されたことを示すものとして合理的に解釈できるようなパフォーマンスを引き出す課題を設計する必要がある。

★「多肢選択式テスト (Multiple-choice tests)」…大人数の生徒の推論を評価するためのデータを

※受験者の推論の証拠は含まれていないため、教師は生徒の誤解を解釈し切れないという課題がある。

生成する。そして、テストのスコアを利用して、教育と学習の有効性についての判断を伝える。

★ニューヨーク州教育局が1999年に開発した「テストのサンプルドラフト」は、多肢選択式問題について、「コア・カリキュラム」のリストからサンプルし、グラフ・漫画・地図・図表の解釈と分析に応用する能力を評価するよう設計されるものだと説明している。

### 2. 事例となる研究の詳細

●目的：多肢選択式歴史問題が引き出すパフォーマンスの種類についての議論に役立つ証拠収集

●方法：①思考発話 (think aloud) →研究参加者に課題遂行時の自らの考えを声に出してもらう

②テスト後のインタビュー→特定の答えを選んだプロセスについて説明してもらう

→大規模な都市部にある小規模で多様性のある学校を選択

→10学年の世界史と地理の教師が第一次世界大戦の終わりから冷戦の始まりまでをカバーする单元について、1ヶ月観察

→同单元の「コア・カリキュラム」のリストに基づく多肢選択式テスト (計15問) を作成し、13人の生徒に思考発話しながら解答してもらう…13人×15問=195のデータが生成された

→同様の13人の生徒に事後インタビューを実施する…①思考発話とインタビューの記録にある記述とテスト問題設計の意図を比較、②生徒が推論に持ち込んだ事実・概念・物語を質的に分析

### 3. 分析結果

- 195 の基礎データを 3 グループに分類：(1)期待と一致、(2)期待と不一致、(3)判定不可能
  - (1)期待と一致：正解の選択肢が選ばれ、受験者が試験内容について知識を持っている証拠がある、不正解の選択肢が選ばれ、受験者が試験内容について知識を持っている証拠がない
  - (2)期待と不一致：解答の正しさがテスト実施者の内容に関する知識の証拠と一致しない
  - (3)判定不可能：10 件あった
- 受験者は、「知識内容 (content knowledge)」、「リテラシー (literacy)」、「テスト習れ (test-wisness)」の 3 つ領域の知識とコンピテンシーを応用し多肢選択式問題を解いていた。

【表 14-1：受験者の知識の証拠採用率と正答率の比較】

多肢選択解答	知識の証拠あり	知識の証拠なし	合計
正解	38.9%	17.3%	56.2%
不正解	8.6%	35.1%	43.7%
合計	47.5%	52.4%	99.9%

(グループ(1)：74%→オレンジ, グループ(2)：25.9%→緑)

#### ● 内容知識：歴史的事実的知識と物語的知識

→ 15 問中、4 問しか正答できなかったクロードと、9 問も正答できたフランクリンの思考発話&インタビューの事例分析に基づく、事実的知識は正解を選択するための前提条件ではなかったが、重要な要素ではあるということが示唆される。また、フランクリンのように、自分の知識に自信を持っているときは、明確な推論をしない可能性が示唆される。

#### 【事例：クロードとフランクリンが解いた設問 1】

設問 1：第二次世界大戦後のニュルンベルク裁判の結果…

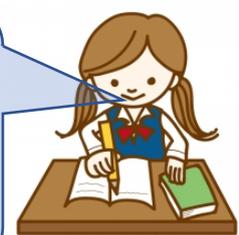
- a) ドイツは 4 つの占領地域に分割された
- b) 国際連合が大量虐殺行為を防止するために結成された
- c) 北大西洋条約機構が共産主義の拡大を阻止するために設立された
- d) ナチスの政治・軍事指導者が行動の責任を問われた (※正答)



(クロード：a)

ニュルンベルク裁判、ニュルンベルク、ドイツ語のように聞こえるなあ。と思うんだけど、どうかな。占領下の 4 分割... 軍事指導者... よくわからないなあ、くそ。難しいな、全部思い出せないぞ。結果は... 結果、結果だ！ドイツは分割。

ナチスの政治・軍事指導者は戦後にニュルンベルク裁判があったので責任を問われたと思う。彼らの中には有罪になる人もいるだろうと。戦犯を裁くことは可能だった。



(フランクリン：d)

#### ● リテラシー：歴史で使用される言語への知的性向 (dispositional orientation)

→ ローマンとフランシーヌの「宥和 (appeasement) 政策」に対する思考発話&インタビューの事例分析に基づく、歴史の言語を可鍛性のある道具として利用しているか、そして道具を使おうとするエージェンシー感を持っているかどうかは歴史的文章の解釈に内容理解を適用できる人 (ローマン) とできない人 (フランシーヌ) を区別する鍵になっている可能性が示唆される。

【事例：ローマンとフランシーヌが解いた設問12】

設問12：宥和政策の最も明確な例は、A～Dのどの発言でしょうか？

A：「戦場で実際に何が起きていたかは、全て秘密だ。大東和共栄圏は、後進国の民族にとって非常に重要だ。」

B：「我々ナチスは、外交政策の目的、すなわちドイツ国民に権利のある土地と土壌を確保しなければならぬ。」

C：「ミュンヘン協定はチェコスロバキアを破壊から救い、ヨーロッパをアルマゲドンから救った。」

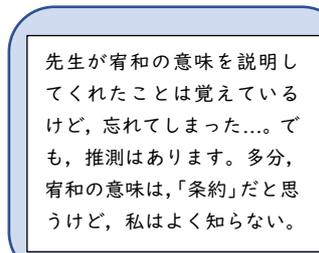
(※正答)

D：「我々は、この島を守らなければならない。どんな犠牲を払ってでもだ。我々は浜辺で戦い、上陸地で戦い、野原で戦い、町の通りでも戦う。我々は決して降伏しない。」



宥和？何だろう？(一旦飛ばす...) えーと。例は、喜ばしいことだからCと言いたいところだな。BやDには喜ばしいことは書いていない。だからCを選んで、それについて考えていたんだ。「宥和」は「快樂」だと思ったからね。

(ローマン：C)



先生が宥和の意味を説明してくれたことは覚えているけど、忘れてしまった...でも、推測はあります。多分、宥和の意味は、「条約」だと思うけど、私はよく知らない。

(フランシーヌ：?)

●テスト習れ (Test-Wisness)：測定されている特性とは論理的に独立している戦略的知識

→受験者の中には、テストのスコアを上げるために使用するヒューリスティックなテストリテラシー(テストを解く際に無意識に利用する経験則やプロセスや態度)を活用している者がいた。経験豊富な受験者(ローレンツ)の思考発話&インタビューの事例分析に基づく、後段の問題の情報が関連する前段の問題の解答を検証したり、疑問を投げかけたり、別の可能性を示唆したりしていないかどうかも見ている可能性が示唆される。

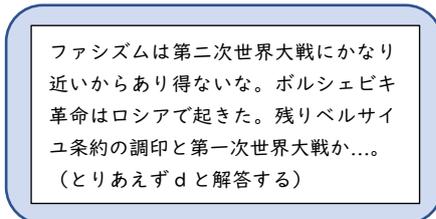
【事例：ローレンツが解いた設問8と10】

設問8：どの出来事が最初に起こった？

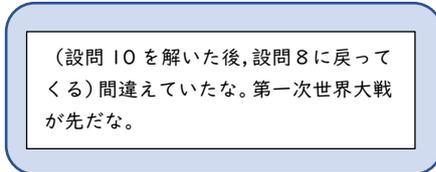
- a) 欧州におけるファシズムの台頭
- b) ボリシェビキ革命
- c) 第一次世界大戦 (※正答)
- d) ベルサイユ条約

設問10：ベルサイユ条約は、第一次世界大戦でのドイツをどう罰した？

- a) ドイツの戦争責任と賠償金 (※正答)
- b) ドイツの4分割
- c) 国際連合による経済制裁の支持
- d) バルカン半島とスペインのドイツ領を剥奪



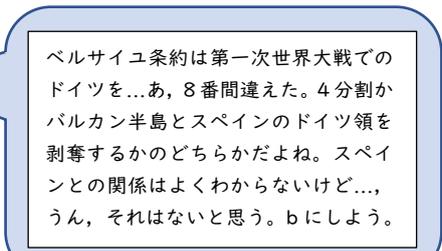
ファシズムは第二次世界大戦にかなり近いからあり得ないな。ボルシェビキ革命はロシアで起きた。残りベルサイユ条約の調印と第一次世界大戦か... (とりあえずdと解答する)



(設問10を解いた後、設問8に戻ってくる)間違えていたな。第一次世界大戦が先だな。



(ローレンツ)



ベルサイユ条約は第一次世界大戦でのドイツを...あ、8番間違えた。4分割かバルカン半島とスペインのドイツ領を剥奪するかのどちらかだよな。スペインとの関係はよくわからないけど...うん、それはないと思う。bにしよう。