

**科学研究費助成事業 研究成果報告書**

平成 29 年 6 月 5 日現在

機関番号：62601

研究種目：基盤研究(C) (一般)

研究期間：2013～2016

課題番号：25381060

研究課題名(和文)カリキュラムの被覆状況に着目したTIMSS理科調査のテスト項目の比較分析

研究課題名(英文)Comparisons of TIMSS science test items based on curriculum coverage

研究代表者

萩原 康仁(Hagiwara, Yasuhito)

国立教育政策研究所・教育課程研究センター基礎研究部・総括研究官

研究者番号：30373187

交付決定額(研究期間全体)：(直接経費) 3,400,000円

研究成果の概要(和文)：TIMSS2011調査の中学校第2学年の物理・化学領域のテスト項目をオーストラリア・イングランド・シンガポールの間で比較した。その結果、知識領域や応用領域において、自国のカリキュラムに含まれないとされた項目群では、能力特性の影響を統制した上で相対的に解きにくくなるという傾向が見られた。この結果は、先行研究ともおおむね整合的であった。

また、TIMSS2007調査とTIMSS2011調査の日本の同学年における生物領域の共通項目の統計的な特性の変化を分析した。この結果、「細胞とその機能」において、平成22年度に中学校第2学年だった生徒の方が能力特性の水準が同じでもやや解きやすい傾向が見られた。

研究成果の概要(英文)：We compared TIMSS 2011 test items of physics and chemistry at the eighth-grade level in Australia, England and Singapore. Some relationships were found between the differences in curriculum coverage in each country and their strengths and/or weaknesses in the cognitive domains of knowing and applying after controlling for the latent traits. These findings were basically consistent with the results of our previous study. We also examined the changes in the statistical characteristics of the common test items of Japanese eighth-graders in biology between TIMSS 2007 and TIMSS 2011. The result showed that, after controlling for the latent trait, the topic area of "Cells and Their Functions" seemed to be easier for the eighth-grade students of 2010-2011 school years than for those of the 2006-2007 school years.

研究分野：教育評価・測定

キーワード：教育評価 カリキュラム 項目分析 TIMSS

### 1. 研究開始当初の背景

国際教育到達度評価学会(以下、IEA)の国際数学・理科教育動向調査(以下、TIMSS調査)等の国際学力調査の結果において、報道等からも明らかのように、一般にとつての主な関心事はテスト得点の国別の順位にある。また、この得点の高低の要因を分析するために、例えば日本では2007年3月に実施されたTIMSS2007調査においても、学校や教師あるいは児童生徒についての様々な背景変数に対する各回答の割合とそれに対応したテスト得点の平均値を国別に集計した表が国際報告書に掲載されており、国内報告書においてもこうした集計表のうち幾つかが取り上げられている。

このTIMSS調査の特徴は小学校第4学年及び中学校第2学年における算数・数学及び理科の教育到達度を評価していることである。このことを踏まえれば、個々のテスト項目で問われた事項を各学校で習っているかどうか、あるいはその前提となる、個々の項目で問われた事項がその国の調査実施学年までのカリキュラムで習うとされているものであるかどうか、というカリキュラムの被覆状況がTIMSS調査の結果に直接影響する要因であると考えことは合理的である。だが、各項目に対する各国におけるカリキュラムの被覆状況の分析(以下、TCMA)の結果は上述の国際報告書にはあるものの付録の扱いであり、上述の国内報告書でも一部が述べられているだけである。

しかしながら、この付録にあるTCMAの結果は、TIMSS2007調査の日本の中学校第2学年理科で特徴的であった。具体的には、この調査における日本の中学校第2学年理科については、カリキュラムに含まれている項目ならば得意であったものの、その数は少なかったということが示唆され、同時に、カリキュラムに含まれていない項目は解きにくかったということも示唆される。

このような調査の特徴とTCMAの結果を踏まえれば、TIMSS調査においては、各国の得点を単に順位づけたり、その要因分析として学校や教師及び児童生徒の背景を取り上げたりすることに先立ち、国等の集団間を比較する際的前提として、カリキュラムの被覆状況の違いを踏まえることが重要であると言える。ただし、TCMAで使用される自国のカリキュラムに含まれた項目群のみを対象にした平均正答率を国や地域間で直接比較することは難しい。さらに、この平均正答率は内容領域や認知的領域を込みにしたものであるため、例えば中学校2学年理科なら物理・化学・生物・地学からなる内容領域のいずれにおいて、あるいは知識・応用・推論からなる認知的領域のいずれにおいて、カリキュラムの被覆状況の影響が見られるのかについても情報を得ることはできない。

そこで、本研究課題ではTIMSS理科調査において、分析対象とした異なる国や調査サイ

クルのそれぞれのカリキュラムに各テスト項目で問われた事項が含まれているかどうかによって、統計的に見た各項目の解きやすさに違いが生じるかどうかを検証しようとし、このことに関わる諸課題も研究の対象としようとした。

### 2. 研究の目的

本研究課題では、国際的な学力調査の一つであるTIMSS調査の理科において、分析対象とした異なる国ないしは調査サイクルのそれぞれのカリキュラムに各テスト項目で問われた事項が含まれているかどうかによって、統計的に見た各項目の解きやすさに違いが生じるかどうかを検証する(なお、1時点における一か国のデータも試行的に分析する)。この検証に当たっては、物理・化学・生物・地学といった内容領域のうちいずれにおいて、あるいは知ること(知識)・応用すること(応用)・推論すること(推論)の認知的領域のうちいずれにおいて、この違いが生じやすいのかを主に調べる。こうした検証によって、我が国における理科の学習指導要領を改善するための実証に基づいた一資料を作成することを目指す。

### 3. 研究の方法

(1) TIMSS2007調査を用いた日本と韓国の中学校第2学年の物理・化学領域の比較

当初予備的に進めていた、TIMSS2007調査を用いた日本と韓国の中学校第2学年の物理・化学領域の比較を行った。統計分析として、項目反応理論(以下、IRT)を用いた特異項目機能(以下、DIF)の分析を行った。

(2) 調査言語が共通の複数の国間におけるTIMSS2011調査の中学校第2学年の物理・化学領域の比較

上記(1)では、調査に使用された言語が異なっていること、及びある調査サイクル(TIMSS2007調査)のデータによるものであるという課題があった。そこで、調査言語がいずれも英語である複数の国(オーストラリア・イングランド・シンガポール)間における、別の調査サイクル(TIMSS2011調査)の中学校第2学年の物理・化学領域の比較を行った。統計分析として、(1)と同様に、IRTを用いたDIFの分析を行った。

(3) 日本のTIMSS2007調査とTIMSS2011調査の中学校第2学年における、生物領域の共通項目の特徴の変化

上記(1)と(2)は、各テスト項目の統計的な特徴である項目特性の違いを、一つの調査サイクルの中で国間の比較で捉えたものである。ここでは、分析対象の国を日本にした上で、複数の調査サイクル間での共通項目の項目特性の違いを比較した。具体的には、

TIMSS2007 調査と TIMSS2011 調査の中学校第 2 学年日本における、生物領域の共通項目の変化を項目母数のずれとして統計的に捉えた。統計分析として、IRT を用いた Item Parameter Drift (以下、IPD) の分析を行った。また、共通項目を生物領域の(単元のような下位領域である)トピックによって質的に分類しておき、統計分析の結果と対応させた。

(4) 日本の TIMSS2007 調査と TIMSS2011 調査の中学校第 2 学年における、地学領域の共通項目の特徴の変化

(3)と同様のことを地学領域で行ったものである。統計分析として、(3)と同様に、IRT を用いた IPD の分析を行った。また、共通項目を地学領域のトピックによって質的に分類しておき、統計分析の結果と対応させた。

(5) TIMSS2011 調査を用いた、小学校理科の生物領域におけるカリキュラムの被覆状況の国際比較と日本の特徴

TIMSS2011 調査のカリキュラムに関する質問調査では、就学前から第 12 学年にいたるまでの、どの学年のカリキュラムにそれぞれのトピックが含まれているかについて、TIMSS 調査の各国調査責任者に尋ねた問がある。ここでは、生物領域における六つのトピックが、参加 59 の国・地域の小学校第 1 学年から第 6 学年までの各学年のカリキュラムで扱われているかどうかについての回答データに対し、カテゴリカル変数に対する確率的な因子分析モデル(特に CT-CM モデル)を当てはめ、マルコフ連鎖モンテカルロ法を用いたベイズ推定により分析した。

(6) TIMSS2011 調査の中学校理科公開問題(日本語版)の項目特性に文書クラスタとの関連度とカリキュラムの被覆状況が与える影響

日本における TIMSS2011 調査の中学校第 2 学年理科の公開問題のテキストデータについて、項目ごとに単語の頻度を数えた。これに基づいて、それぞれの項目内の各単語の重みを示す tf-idf を要素とした単語-文書行列を作成した。この行列に対して非負値行列因子分解を行い、複数個のクラスタを得ることとした。さらに、各項目の各クラスタとの関連度及びカリキュラムの被覆状況が、難易度や識別力といった生徒の解答データから得られる各項目の統計的な特性の違いを説明するかについて、交差変数効果を持つ IRT モデルを当てはめ、マルコフ連鎖モンテカルロ法を用いたベイズ推定により分析した。

#### 4. 研究成果

(1) TIMSS2007 調査を用いた日本と韓国の中学校第 2 学年の物理・化学領域の比較

分析の結果、特に知識領域や応用領域の項目において、TCMA に基づいたカリキュラムの

被覆状況に関連した、相対的な得意不得意の傾向が見られた。その一方で、推論領域の項目においては TCMA に関連した相対的な得意不得意の傾向は見られなかった。

(2) 調査言語が共通の複数の国間における TIMSS2011 調査の中学校第 2 学年の物理・化学領域の比較

分析の結果、物理・化学の応用領域において、シンガポールの生徒にとって自国のカリキュラムにのみ内容が含まれないとされた項目群では、能力特性の影響を統制した上で相対的に解きにくくなるという傾向が見られた。また、物理の知識領域において、オーストラリアの生徒にとって自国の州のカリキュラムでのみ含まれないとされた項目群では、相対的にやや解きにくくなるという傾向が見られた。これらの結果は(1)の結果とおおむね整合的であり、こうした関連性が一定程度確かなものであることが示唆された。

(3) 日本の TIMSS2007 調査と TIMSS2011 調査の中学校第 2 学年における、生物領域の共通項目の特徴の変化

分析の結果、生物領域の「細胞とその機能」のトピックにおいては、平成 22 年度に中学校第 2 学年だった生徒の方が平成 18 年度に同学年だった生徒に比べて、能力特性の水準が同じだとしてもやや解きやすい傾向が見られた。この分析結果は、内容のつながりの明確化と学習時期の移行といったカリキュラム改訂の効果に関連するものと考察された。

(4) 日本の TIMSS2007 調査と TIMSS2011 調査の中学校第 2 学年における、地学領域の共通項目の特徴の変化

分析の結果、「地球のプロセス、循環と歴史」というトピックにおいては、能力特性の水準が同じだとしても、平成 22 年度に中学校第 2 学年だった生徒の方が、平成 18 年度にこの学年だった生徒よりも解きやすかったという傾向が見られた。しかしながら、(3)で見られたような、生物領域の「細胞とその機能」における学習時期の移行といった中学校学習指導要領間での違いは、「地球のプロセス、循環と歴史」のトピックに関連するところでは見られなかった。

(5) TIMSS2011 調査を用いた、小学校理科の生物領域におけるカリキュラムの被覆状況の国際比較と日本の特徴

分析の結果、生物領域における六つのトピックのうち、「人間の健康」と「環境の変化」について、参加国・地域の平均に比べれば日本の小学校理科カリキュラムで扱われていないことが示唆された。なお、異なる事前分布の条件を複数用意してベイズ推定を行ったが、これら複数の条件下でも上記の結果はおおむね変わらなかった。この結果について、

「人間の健康」のトピックは理科ではなく体育等の他の教科や学校全体の教育活動で扱われていると考えられ、児童のこのトピックについてのテスト項目群の通過率も国際平均に比べ高いままであった。一方で、「環境の変化」のトピックにおいては、児童のこのトピックについてのテスト項目群の多くにおいて通過率が必ずしも高いとは言えず、ESD（持続可能な開発のための教育）との関連を踏まえれば、このトピックに関連する内容をカリキュラムに含める余地が残されていることが示唆された。

(6) TIMSS2011 調査の中学校理科公開問題(日本語版)の項目特性に文書クラスタとの関連度とカリキュラムの被覆状況が与える影響

非負値行列因子分解の結果、四つのクラスタが得られた。この四つのクラスタの解釈から、TIMSS2011 理科の公開問題では、ある語が他の特定の語と共起していることがあり、こうした語のまとまりが理科の内容の一部を指し示すことが示唆された。

IRT 分析の結果、履修している項目の方が易しく識別力は高い傾向にあること、四つのうちあるクラスタとの関連度が大きい項目ほど易しいこと、及び、別のクラスタとの関連度が大きい項目ほど難しく識別力は低いことが示唆された。これら三つの結果のうち、1点目と3点目から、「エネルギー」という語が含まれた、日本のカリキュラムに含まれていない生物領域の二つの項目(共に選択式)の日本の通過率が、日本の選択式の項目に対する平均的な通過率より大幅に低いことを確認した。

本研究課題を通じて、以下のデータベースを用いた。記して謝辞とする。

- ・TIMSS 2007 Assessment. Copyright © 2009 International Association for the Evaluation of Educational Achievement (IEA). Publisher: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- ・TIMSS 2011 Assessment. Copyright © 2013 International Association for the Evaluation of Educational Achievement (IEA). Publisher: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計3件)

Matsubara, K., Hagiwara, Y., & Saruta, Y. (2016). A statistical analysis of the characteristics of the intended curriculum for Japanese

primary science and its relationship to the attained curriculum. *Large-scale Assessments in Education*, 4:13. (査読有)

doi: 10.1186/s40536-016-0028-0

萩原康仁・松原憲治 (2014). オーストラリア, イングランド, シンガポールを対象とした TIMSS2011 調査の第 8 学年物理・化学領域におけるカリキュラムの被覆状況を関連付けた IRT 分析 日本テスト学会誌, 10, 95-113. (査読有)

松原憲治・萩原康仁 (2014). 国際的な学力調査から見る日本の理科の学力とカリキュラムの変化との関連: 近年の TIMSS 調査における中学校第 2 学年生物領域の共通項目の変化に基づいて 理科教育学研究, 55, 69-80. (査読有)

doi: 10.11639/sjst.sp13015

[学会発表](計 10 件)

萩原康仁・松原憲治 TIMSS2011 中学校理科公開問題(日本語版)の項目特性に文書クラスタとの関連度と履修状況が与える影響の分析 日本教育心理学会第 58 回総会 2016 年 10 月 8 日 サンポートホール高松(香川県高松市)

萩原康仁・松原憲治 加法モデルを用いた TIMSS カリキュラム質問紙の分析: シミュレーションと実データの分析 日本テスト学会第 13 回大会 2015 年 9 月 10 日 関西大学(大阪府吹田市)

松原憲治・萩原康仁・猿田祐嗣 TIMSS2011 の結果を用いた日本の小学校理科カリキュラムの分析: カリキュラム質問紙のデータから 日本理科教育学会第 65 回全国大会 2015 年 8 月 1 日 京都教育大学(京都市)

Hagiwara, Y., & Matsubara, K. The relationship between the statistical changes in common items and the revision of the national curriculum in Japan: A (constrained) item parameter drift analysis of TIMSS 2007 and TIMSS 2011 assessments of eighth graders on earth science. The 6th IEA International Research Conference. Jun. 25, 2015. Cape Town, South Africa.

Matsubara, K., & Hagiwara, Y. A statistical analysis of the characteristics of the intended curriculum of Japanese primary science and its relationships with the attained curriculum. The 6th IEA International Research Conference. Jun. 25, 2015. Cape Town, South Africa.

萩原康仁・松原憲治 近年の TIMSS 調査における中学校第 2 学年地学領域の共通項目の統計的な特徴の変化とトピックと

の関連 日本教育心理学会第 56 回総会  
2014 年 11 月 07 日 神戸国際会議場(神戸市)

萩原康仁・松原憲治 カリキュラムの被  
覆状況を関連付けたオーストラリアとニ  
ュージーランドにおける TIMSS2011 調査  
第 8 学年物理領域の比較分析 日本理科  
教育学会第 64 回全国大会 2014 年 8 月  
23 日 愛媛大学(愛媛県松山市)

萩原康仁・松原憲治 日本における中学  
校第 2 学年理科の共通項目を対象とした  
TIMSS2011 調査と TIMSS2007 調査の比較  
分析 日本テスト学会第 11 回大会  
2013 年 8 月 28 日 九州大学(福岡市)

萩原康仁・松原憲治 TIMSS2011 調査(理  
科)の結果と分析(2): 項目反応理論を  
用いた日本の中学生における TIMSS2007  
調査との比較分析 日本理科教育学会第  
63 回全国大会 2013 年 8 月 10 日 北海  
道大学(札幌市)

Hagiwara, Y., & Matsubara, K. A DIF  
analysis of TIMSS-2007 assessment in  
physics and chemistry focusing on the  
Matching of the Test Items and the  
Curricula: The Comparison of Japanese  
and Korean Eighth Graders. The 5th  
IEA International Research Conference.  
Jun. 26, 2013. Singapore.

## 6. 研究組織

### (1) 研究代表者

萩原 康仁 (HAGIWARA, Yasuhito)  
国立教育政策研究所・教育課程研究センタ  
ー基礎研究部・総括研究官  
研究者番号: 3 0 3 7 3 1 8 7

### (2) 研究分担者

松原 憲治 (MATSUBARA, Kenji)  
国立教育政策研究所・教育課程研究センタ  
ー基礎研究部・総括研究官  
研究者番号: 1 0 5 4 9 3 7 2