

TALIS2013における日本の教員の自己効力感についての確認的な
多次元項目反応モデルに基づいた一検討An investigation of Japanese teachers' efficacy in TALIS 2013 based
on confirmatory multidimensional item response models

萩原 康仁*・松原 憲治**

HAGIWARA Yasuhito and MATSUBARA Kenji

Abstract

To examine the self-efficacy of Japanese teachers in TALIS 2013, it was important to analyze the relationships between the individual differences in each of the three sub-scales of the self-efficacy and their responses to each questionnaire item, regarding response categories as ordinal. The purpose of this study was to investigate the tendency of the responses of Japanese secondary school teachers for each level of the respective sub-scales. The authors applied confirmatory multidimensional ordinal item response models to the data on the self-efficacy of Japanese teachers in TALIS 2013.

The results showed that, the average teachers tended to respond “To some extent”, which was the second least positive response category of the four, to the respective items on the efficacy in student engagement. To the items on efficacy in classroom management, the average teachers tended to respond to about the same degree “To some extent” or “Quite a bit”, which was the second most positive category. The results also revealed that, for all items on the three aspects of efficacy, the teachers with one standard deviation below the mean tended to respond “To some extent”. As for the efficacy in student engagement, on the other hand, there were items that the teachers with one standard deviation above the mean did not tend to respond “Quite a bit.” We obtained some suggestions about the professional development of teachers on self-efficacy.

* 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
** 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

1. 問題

教員が自信を持って教職に当たることができているということは、児童生徒にとっても重要なことである。しかしながら、OECDによる国際教員指導環境調査の2013年調査（以下、TALIS2013とする）の結果によると、表1から表3に示した教員の自己効力感に関する三つの側面とそれぞれの側面で尺度を構成する計12項目⁽¹⁾の回答結果から、“日本では自己効力感の高い教員の割合が16%から54%と低い”（国立教育政策研究所, 2014, p.191）という実態が指摘された。さらに、“OECD加盟国の平均と比較して低いだけでなく、TALIS2013年調査参加国の中で最も低い”（国立教育政策研究所, 2014, p.191）とも注記された。

表1 「学級運営」を構成する項目群

問34	項目文
(4)	学級内の秩序を乱す行動を抑える
(6)	自分が生徒にどのような態度・行動を期待しているか明確に示す
(8)	生徒を教室のきまりに従わせる
(9)	秩序を乱す又は騒々しい生徒を落ち付かせる

国立教育政策研究所（2014, P227）より作成

表2 「教科指導」を構成する項目群

問34	項目文
(3)	生徒のために発問を工夫する
(10)	多様な評価法を活用する
(11)	生徒がわからない時には、別の説明の仕方を工夫する
(12)	様々な指導方法を用いて授業を行う

国立教育政策研究所（2014, P227）より作成

表3 「生徒の主体的学習参加の促進」を構成する項目群

問34	項目文
(1)	生徒に勉強ができると自信を持たせる
(2)	生徒が学習の価値を見いだせるよう手助けする
(5)	勉強にあまり関心を示さない生徒に動機付けをする
(7)	生徒の批判的思考を促す

国立教育政策研究所（2014, P227）より作成

TALIS2013は、学校の学習環境と教員の勤務環境に焦点を当てた、通常の仕事として指導を行う教員及び校長に対して実施された質問紙調査であり、“職能開発などの教員の環境、学校での指導状況、教員への評価やフィードバックなどについて、国際比較可能なデータを収集し、教育に関する

分析や教育政策の検討に資することを目指し”（国立教育政策研究所，2014，p.5）たものである。TALIS2013の結果は平成26年6月25日に公表され、このうち、教員の自己効力感については、上記のような低い実態が指摘されたところである。この指摘は、OECD（2014a）において、12項目のそれぞれにおける回答カテゴリ（いずれも、「非常に良くできている」、「かなりできている」、「ある程度できている」、「まったくできていない」の四つからなっていた）のうち、前者二つ（「非常に良くできている」、「かなりできている」）のいずれかに反応した教員の割合の計を当該項目における自己効力感の高い教員の割合であるとして集計されたことに基づいて行われた。また、OECD（2014a，p.430）のリンク先（<http://dx.doi.org/10.1787/888933047482>）には、各項目の回答カテゴリごとの反応率について記載されている。この情報に基づいて、“日本の教員は、いずれの質問項目についても「ある程度できている」と回答した割合が最も高い”（国立教育政策研究所，2014，p.191）ことが分かり、日本の教員はそれぞれの自己効力感に関する項目について「まったくできていない」という実態ではないことが伺える。さらに、日本の教員の回答について特徴的な点に着目すると、三つの側面のうち「学級運営」と「教科指導」に関する項目については高い自己効力感を持つ割合が比較的高い（一項目を除き、43%から54%）一方で、「生徒の主体的学習参加の促進」に関する項目について高い自己効力感を持つ割合は16%から26%と特に低いことが指摘されている（国立教育政策研究所，2014）。以上のように、日本では自己効力感の高い教員の割合が高くないという課題がある。

この課題を改善するための方策として、例えば自己効力感の向上に資する研修を教員に実施することが考えられる。この研修を計画する際には、自己効力感のどの側面において実施しやすいかについて、教員の個人差を関連させた上で予測しておくことは重要である。したがって、日本の教員の自己効力感について、個人差に関する実態を側面ごとに捉えるという分析の視点が求められる。しかしながら、上述の分析はいずれも、どのくらいの割合で日本の教員が各項目の回答カテゴリに反応しているかという反応率の集計であり、教員の自己効力感の個人差に着目した分析の視点ではない。すなわち、各項目の反応率から、平均的な自己効力感の教員、あるいは相対的に高い（若しくは低い）自己効力感の教員が、各項目に対してどの回答カテゴリに反応しやすいか、あるいはしにくいかを、自己効力感の側面ごとに捉えることは難しいということである。したがって本研究では、日本において平均的な自己効力感の教員のみならず、相対的に高い教員や低い教員が、自己効力感の各側面でどの回答カテゴリに反応する傾向にあるのかを検討する。

一方で、教員の自己効力感における三つの側面の個人差を分析する方法として、TALIS2013のテクニカルレポート（OECD，2014b，p.200）では、各側面を下位尺度としてそれぞれの因子に対応付け、各項目への回答を連続変数とみなした3因子の確認的因子分析モデルの推定を行っている。ただし、この分析方法では、本来は順序尺度である回答カテゴリを間隔尺度と扱っており、自己効力感の水準ごとに各項目の回答カテゴリごとの反応傾向を捉えるのは難しい。

以上の議論から、TALIS2013に基づいて日本の教員の自己効力感について検討する際には、回答カテゴリを本来の順序尺度として扱った上で、各回答カテゴリの反応傾向を自己効力感の各側面の個人差との関係でモデル化した分析に基づくことが望まれる。このため、本研究では、TALIS2013の公開データのうち、教員の自己効力感に関する日本の回答データに対して、回答を順序尺度として扱った確認的な多次元項目反応モデルを適用する。このモデルを用いることによって、教員の自己効力感の個人差を側面ごとに潜在特性値として表すことができ、各項目における各回答カテゴリへの反応確率を、対応する自己効力感の側面における潜在特性値の水準ごとに推定することが可能となる。したがって、平均的な自己効力感の教員、あるいは相対的に高い（若しくは低い）教員が、

各項目に対してどの回答カテゴリに反応しやすいか、あるいはしにくいかについて、自己効力感の側面ごとに捉えることができる。

2. 目的

本研究では、日本の前期中等教育段階における教員が自己効力感の各側面でどの回答カテゴリに反応する傾向にあるのかについて、自己効力感の個人差との関連で検証することを目的とする。このことによって、日本の教員がどのカテゴリに回答する傾向にあるのかについて、自己効力感の各側面の水準ごとに捉えることができる。

さらに、教員の自己効力感の向上に関する研修への示唆を、その個人差との関連で得る。

3. 方法

3.1. 分析対象

TALIS2013 の公開データ (http://stats.oecd.org/Index.aspx?datasetcode=talis_2013%20) にある日本の 192 校の中学校及び中等教育学校の前期課程における 3,484 名の教員を対象とした。分析対象項目は、自己効力感に関する 12 項目 (問 34 の(1)~(12)) であった。これらの項目全てに対して無回答であった 21 名を分析対象から除外した。最終的な分析対象者は 192 校の 3,463 名の教員とした。

表 1 から表 3 にある通り、(4)、(6)、(8)、(9)の 4 項目は「学級運営」を、(3)、(10)、(11)、(12)の 4 項目は「教科指導」を、(1)、(2)、(5)、(7)の 4 項目は「生徒の主体的学習参加の促進」を測定しているものとされた (国立教育政策研究所、2014)。なお、日本のデータに対して上述の三つの因子と各因子から各項目へのパスの配置を仮定した 3 因子の確認的因子分析を適用した結果が OECD (2014b) に記載されており、それによると、適合度指標は CFI = 0.957、TLI = 0.944、RMSEA = 0.054、SRMR = 0.047 であり、他の参加各国において同じモデルを当てはめた際の適合度指標も一覽で見ても概して許容できる適合の程度とされた。

3.2. 分析モデル

本研究では、後述する二つの確認的な多次元項目反応モデルを用いた。具体的には、項目に対する反応の形式として、多値型の項目反応モデルである段階反応モデル (Samejima, 1969) 及び一般化部分採点モデル (Muraki, 1992) の二つのモデルに基づいた。それぞれのモデルについて、潜在特性の数を 3 とし、各潜在特性から各項目へのパスの配置について、表 1 から表 3 に示した尺度と項目との対応関係に基づいた確認的な多次元項目反応モデルを適用した⁽²⁾。このため、ある項目に影響を与える潜在特性は、どの項目でも三つのうちいずれか一つであるとした。加えて、それぞれのモデルにおいて、潜在特性について平均 0、分散 1 の多変量正規分布を仮定し、潜在特性間の相関を仮定した。

3.3. 確認的な多次元段階反応モデル

潜在特性の数が一つである 1 次元の段階反応モデル (Samejima, 1969) を、複数の潜在特性を仮定するように拡張した多次元段階反応モデル (Muraki & Carlson, 1995) がある。本研究では、後述するソフトウェアで用いられているロジスティックモデルで表現されたモデル式 (Asparouhov &

Muthén, 2015) に基づいた。具体的には、潜在特性が \mathbf{f} (本研究では 3 行 1 列の縦ベクトル) である際に、 i 番目の項目に対してカテゴリ j の回答をする確率 $P_{ij}(\mathbf{f})$ を、 j が最初のカテゴリ (本研究では、最も否定的な回答カテゴリ) の場合は

$$P_{ij}(\mathbf{f}) = \frac{1}{1 + \exp(-\tau_{ij} + \boldsymbol{\lambda}'_i \mathbf{f})}$$

とし、 j が最後のカテゴリ (最も肯定的な回答カテゴリ) の場合は

$$P_{ij}(\mathbf{f}) = 1 - \frac{1}{1 + \exp(-\tau_{ij-1} + \boldsymbol{\lambda}'_i \mathbf{f})}$$

とし、 j が中間のカテゴリの場合は

$$P_{ij}(\mathbf{f}) = \frac{1}{1 + \exp(-\tau_{ij} + \boldsymbol{\lambda}'_i \mathbf{f})} - \frac{1}{1 + \exp(-\tau_{ij-1} + \boldsymbol{\lambda}'_i \mathbf{f})}$$

とするモデルである。 τ_{ij} は項目 i の当該カテゴリまでのしきい値を示し、項目あたり $j-1$ (本研究では 3) 個推定される。また、 $\boldsymbol{\lambda}_i$ は項目 i の因子負荷量を示す縦ベクトルであり、本研究では例えば、

$$\boldsymbol{\lambda}'_1 = (0 \quad 0 \quad \lambda_{13})$$

$$\boldsymbol{\lambda}'_3 = (0 \quad \lambda_{32} \quad 0)$$

$$\boldsymbol{\lambda}'_4 = (\lambda_{41} \quad 0 \quad 0)$$

のように、各項目は対応するいずれか一つの潜在特性にのみ影響を受けるという確認的なモデルを仮定する。

3.4. 確認的な多次元一般化部分採点モデル

潜在特性の数が一つである 1 次元の一般化部分採点モデル (Muraki, 1992) を、複数の潜在特性を仮定するように拡張した多次元一般化部分採点モデル (Yao & Schwarz, 2006) がある。本研究では、後述するソフトウェアで用いられているモデル式 (Asparouhov & Muthén, 2015) に基づいた。具体的には、潜在特性が $\boldsymbol{\eta}$ (本研究では 3 行 1 列の縦ベクトル) である際に、 i 番目の項目に対してカテゴリ k ($k=0, \dots, m-1$) の回答をする確率 $P_{ik}(\boldsymbol{\eta})$ を

$$P_{ik}(\boldsymbol{\eta}) = \frac{\exp(\sum_{h=0}^k (\boldsymbol{\beta}'_i \boldsymbol{\eta} - \tau_{h,i}))}{\sum_{g=0}^{m-1} \exp(\sum_{h=0}^g (\boldsymbol{\beta}'_i \boldsymbol{\eta} - \tau_{h,i}))}$$

とするモデルである。 $\tau_{h,i}$ は項目 i の隣接するカテゴリ間のしきい値を示す。各項目で、通常、 $\tau_{0,i} = 0$ と制約を置くため、項目あたり $m-1$ (本研究では 3) 個推定される。また、 $\boldsymbol{\beta}_i$ は項目 i の因子負荷量を示す縦ベクトルであり、本研究では先と同様に、

$$\boldsymbol{\beta}'_1 = (0 \quad 0 \quad \beta_{13})$$

$$\boldsymbol{\beta}'_3 = (0 \quad \beta_{32} \quad 0)$$

$$\boldsymbol{\beta}'_4 = (\beta_{41} \quad 0 \quad 0)$$

のように、各項目は対応するいずれか一つの潜在特性にのみ影響を受けるという確認的なモデルを仮定する。

3.5. 分析方法及び分析条件

TALIS2013 の日本の標本抽出法は、学校を一次抽出単位とする層化二段抽出法であった。このため、学校 ID によるクラスタリングと各教師の標本加重を考慮する分析を行った。ただし、本来は抽出に用いた層の情報も分析の際に考慮すべきではあるが、公開データになく使用できなかった。分析には Mplus version 7.4 (Muthén & Muthén, 1998–2015) というソフトウェアを用いた。このソフトウェアで MLR (maximum likelihood estimator with robust standard errors using a numerical integration algorithm) と呼ばれる推定法を用い、欠測については Missing at random (Rubin, 1976) を仮定した。どちらのモデルの推定でも、潜在変数あたりの求積点の数は 20 とした。

4. 結果

4.1. 両モデルの比較

情報量規準の観点からは、確認的な多次元段階反応モデル (AIC = 60596.691、BIC = 60910.335) の方が確認的な多次元一般化部分採点モデル (AIC = 60688.603、BIC = 61002.248) よりも良かった³⁾。ただし本研究では、考察のため両モデルの結果を示す。

4.2. 確認的な多次元段階反応モデルの結果

確認的な多次元段階反応モデルにおける母数の推定値と標準誤差を表 4 に示す。また、各項目について、対応する潜在特性の各水準における項目反応カテゴリ特性曲線を図 1 から図 12 に示す⁴⁾。

表 4 確認的な多次元段階反応モデルの推定値と標準誤差

自己効力感	項目番号	λ_{i1}	λ_{i2}	λ_{i3}	τ_{i1}	τ_{i2}	τ_{i3}
学 級 運 営 (A)	問34(4)	3.331 (0.166)			-7.861 (0.342)	<i>-0.250</i> (0.094)	4.417 (0.197)
	問34(6)	1.677 (0.076)			-5.323 (0.177)	-0.189 (0.055)	3.107 (0.099)
	問34(8)	3.885 (0.233)			-9.714 (0.596)	<i>0.122</i> (0.121)	5.660 (0.286)
	問34(9)	4.631 (0.267)			-9.988 (0.594)	<i>0.028</i> (0.140)	6.239 (0.322)
教 科 指 導 (B)	問34(3)		2.091 (0.091)		-5.785 (0.214)	0.469 (0.066)	4.428 (0.152)
	問34(10)		1.986 (0.082)		-3.608 (0.123)	1.573 (0.072)	4.961 (0.147)
	問34(11)		2.933 (0.118)		-8.080 (0.335)	-0.391 (0.082)	4.326 (0.165)
	問34(12)		3.010 (0.136)		-6.692 (0.258)	0.534 (0.087)	4.876 (0.191)
参 生 加 徒 の の 促 主 進 体 (C 学 習)	問34(1)			2.810 (0.154)	-5.614 (0.253)	3.062 (0.145)	7.293 (0.306)
	問34(2)			3.351 (0.219)	-6.896 (0.357)	2.401 (0.162)	7.917 (0.435)
	問34(5)			2.276 (0.108)	-4.228 (0.157)	2.196 (0.091)	6.114 (0.204)
	問34(7)			1.073 (0.069)	-1.834 (0.072)	1.987 (0.064)	4.600 (0.146)

カッコ内はロバスト標準誤差。斜体以外は $p < .001$ で有意。
AB間の相関 .676 (.016), AC間の相関 .653 (.019), BC間の相関 .791 (.015)。

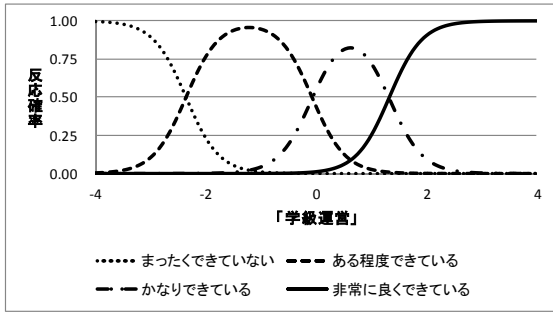


図1 問34(4)の項目反応カテゴリ特性曲線(段階反応モデル)

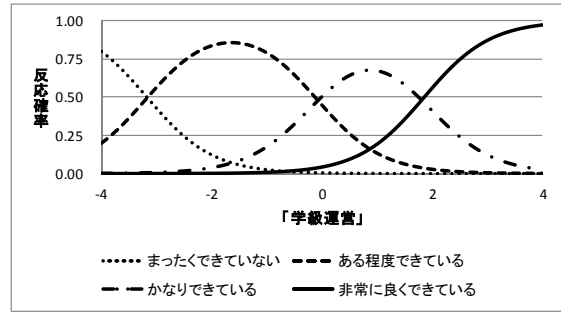


図2 問34(6)の項目反応カテゴリ特性曲線(段階反応モデル)

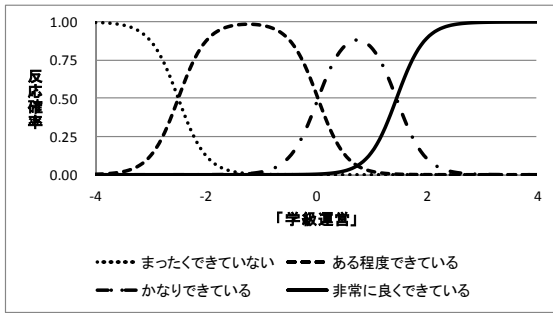


図3 問34(8)の項目反応カテゴリ特性曲線(段階反応モデル)

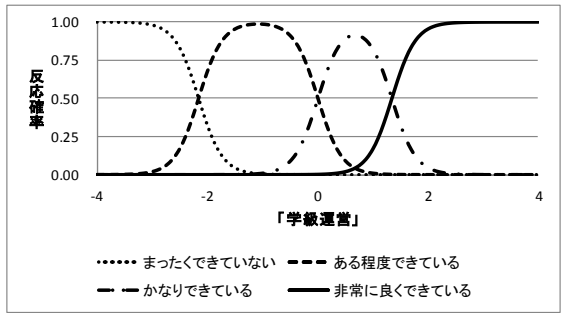


図4 問34(9)の項目反応カテゴリ特性曲線(段階反応モデル)

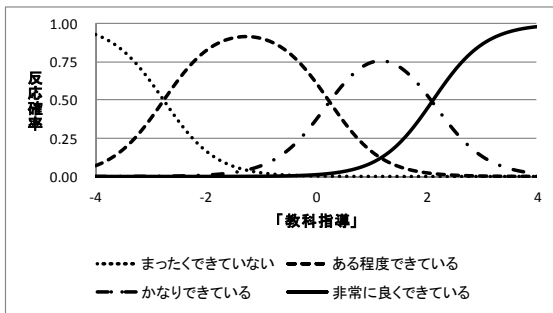


図5 問34(3)の項目反応カテゴリ特性曲線(段階反応モデル)

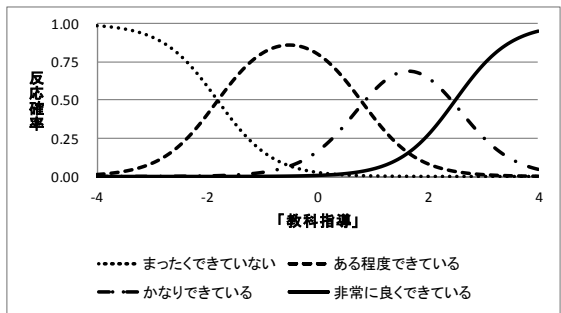


図6 問34(10)の項目反応カテゴリ特性曲線(段階反応モデル)

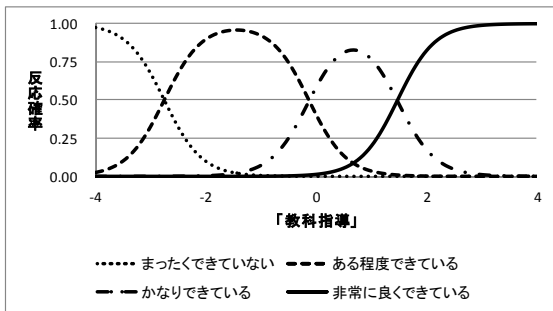


図7 問34(11)の項目反応カテゴリ特性曲線(段階反応モデル)

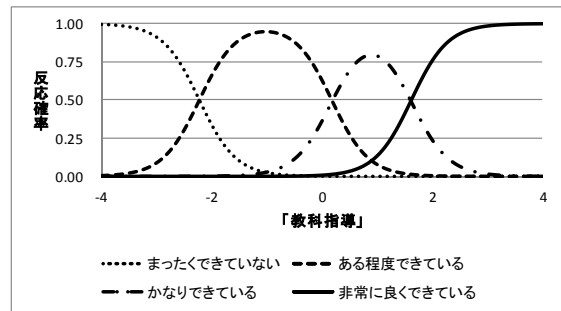


図8 問34(12)の項目反応カテゴリ特性曲線(段階反応モデル)

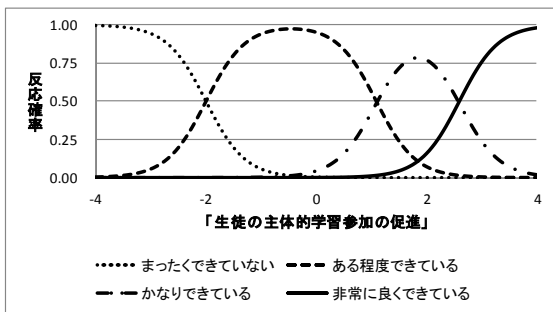


図9 問34(1)の項目反応カテゴリ特性曲線(段階反応モデル)

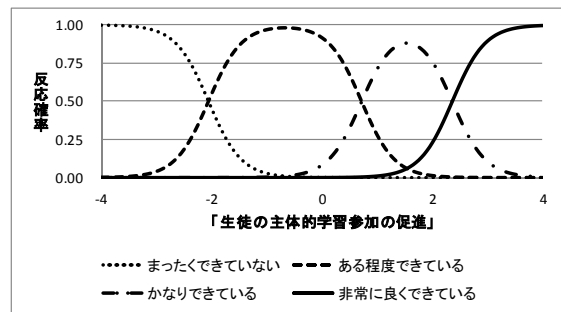


図10 問34(2)の項目反応カテゴリ特性曲線(段階反応モデル)

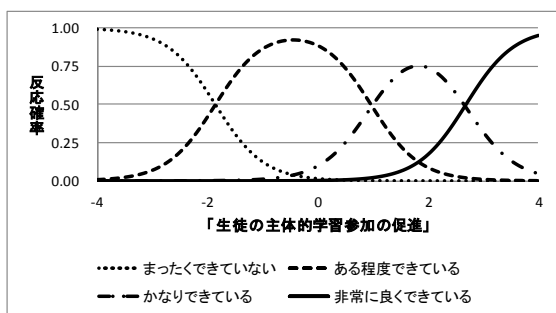


図 11 問 34(5)の項目反応カテゴリ特性曲線 (段階反応モデル)

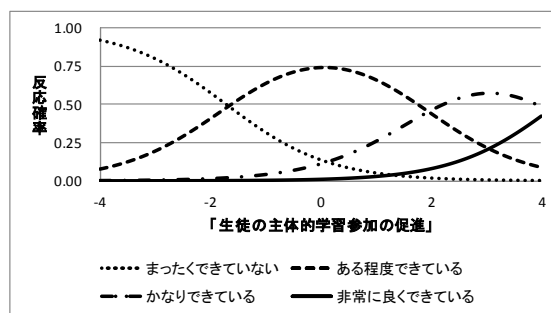


図 12 問 34(7)の項目反応カテゴリ特性曲線 (段階反応モデル)

第一に、平均的な自己効力感の教員として、各側面で潜在特性値が 0 (偏差値換算で 50) を想定した場合の、三つの側面で共通する結果を述べる。表 4 から、いずれの τ_{11} も高度に有意 ($p < .001$) であり、推定値はいずれも負で、絶対値も大きかった。各図においても、「まったくできていない」についての曲線は左寄りであり、潜在特性値が 0 の場合のこの回答カテゴリへの反応確率がいずれもほぼ 0 であった。したがって、日本の前期中等教育段階における平均的な自己効力感の教員は、三つの側面のどの項目でも、最も否定的な「まったくできていない」よりもそれを超えるカテゴリの方により回答する傾向にあった。

また、いずれの τ_{13} も高度に有意 ($p < .001$) であり、推定値はいずれも正で、絶対値も大きかった。各図においても、「非常に良くできている」についての曲線は右寄りであり、潜在特性値が 0 の場合のこの回答カテゴリへの反応確率がいずれもほぼ 0 であった。したがって、日本の前期中等教育段階における平均的な自己効力感の教員は、三つの側面のどの項目でも、最も肯定的な「非常に良くできている」のカテゴリまで回答する傾向にはなかった。

第二に、平均的な自己効力感の教員について、各側面での特徴的な結果を述べる。「学級運営」については、 τ_{12} の絶対値が 0 付近であり、比較的大きい標本サイズであるが有意ではない項目も複数あった。図 1 から図 4 においても、潜在特性値が 0 付近での「まったくできていない」と「ある程度できている」の反応確率の和と、「かなりできている」と「非常に良くできている」の反応確率の和はいずれも 0.5 前後であった。したがって、反応率の集計ではいずれも「ある程度できている」が高いものの、本モデルの推定結果では、「学級運営」について日本の前期中等教育段階における平均的な自己効力感の教員は、「学級運営」を測定しているこれらの項目群において、「ある程度できている」までのカテゴリに回答しているか、より肯定的なカテゴリに回答をしているか分かれる傾向にあった。

また、「生徒の主体的学習参加の促進」については、いずれの τ_{12} も高度に有意 ($p < .001$) であり、推定値はいずれも正で、絶対値も大きかった。図 9 から図 12 において、「非常に良くできている」だけでなく、「かなりできている」についての曲線も右寄りであり、潜在特性値が 0 の場合の「かなりできている」への反応確率は低かった。したがって、「生徒の主体的学習参加の促進」について日本の前期中等教育段階における平均的な自己効力感の教員は、「生徒の主体的学習参加の促進」を測定している項目群において、「ある程度できている」までのカテゴリに回答する傾向にあった。

なお、「教科指導」については、(11)以外は τ_{12} が正であり、高度に有意 ($p < .001$) であった。図 7 以外の図 5、図 6、図 8 において、潜在特性値が 0 の場合の反応確率は、「ある程度できている」が相対的に高かった。

第三に、自己効力感が相対的に高い教員として、各側面で潜在特性値が 1 (偏差値換算で 60) を

想定した場合の結果を述べる。こうした教員が回答するカテゴリの傾向について、「学級運営」と「教科指導」においてはどの項目でも、「かなりできている」と回答する確率が最も高かった。しかしながら、「生徒の主体的学習参加の促進」においては、複数の項目（(1)、(7)）で「かなりできている」ではなく「ある程度できている」への反応確率が最も高かった。(5)では、「ある程度できている」と「かなりできている」への反応確率がそれぞれ約48%と約50%であり、同程度であった。

第四に、自己効力感が相対的に低い教員として、各側面で潜在特性値が-1（偏差値換算で40）を想定した場合の結果を述べる。こうした教員が反応する回答カテゴリの傾向について、三つの側面のどの項目でも、「ある程度できている」と回答する確率が最も高かった。

4.3. 確認的な多次元一般化部分採点モデルの結果

確認的な多次元一般化部分採点モデルにおける母数の推定値と標準誤差を表5に示す。また、各項目について、対応する潜在特性の各水準における項目反応カテゴリ特性曲線を図13から図24に示す⁽⁵⁾。

表5 確認的な多次元一般化部分採点モデルの推定値と標準誤差

自己効力感	項目番号	β_{11}	β_{12}	β_{13}	$\tau_{1,j}$	$\tau_{2,j}$	$\tau_{3,j}$
学級 運営 (A)	問34(4)	3.199 (0.180)			-7.638 (0.366)	<i>-0.198</i> (0.093)	4.200 (0.220)
	問34(6)	1.449 (0.076)			-4.803 (0.186)	<i>-0.086</i> (0.053)	2.606 (0.115)
	問34(8)	3.787 (0.246)			-9.554 (0.618)	<i>0.162</i> (0.119)	5.508 (0.303)
	問34(9)	4.492 (0.262)			-9.774 (0.586)	<i>0.060</i> (0.137)	6.033 (0.321)
教科 指導 (B)	問34(3)		1.937 (0.097)		-5.496 (0.229)	0.503 (0.064)	3.997 (0.178)
	問34(10)		1.753 (0.080)		-3.304 (0.127)	1.544 (0.069)	4.154 (0.174)
	問34(11)		2.833 (0.127)		-7.913 (0.356)	-0.342 (0.082)	4.131 (0.181)
	問34(12)		2.860 (0.145)		-6.474 (0.274)	0.564 (0.084)	4.558 (0.214)
参 加 徒 の の 促 主 進 体 (C 学 習)	問34(1)			2.713 (0.160)	-5.497 (0.261)	3.021 (0.146)	6.784 (0.355)
	問34(2)			3.289 (0.225)	-6.833 (0.369)	2.395 (0.162)	7.641 (0.460)
	問34(5)			2.137 (0.112)	-4.058 (0.164)	2.151 (0.089)	5.472 (0.243)
	問34(7)			0.877 (0.062)	-1.595 (0.071)	1.871 (0.066)	3.039 (0.181)

カッコ内はロバスト標準誤差。斜体以外は $p < .001$ で有意。
AB間の相関 .677 (.016), AC間の相関 .654 (.019), BC間の相関 .791 (.015)。

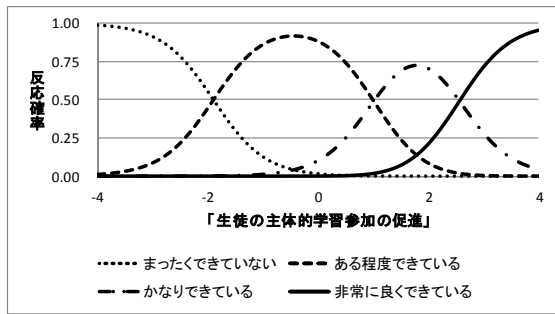


図 23 問 34(5)の項目反応カテゴリ特性曲線（一般化部分採点モデル）

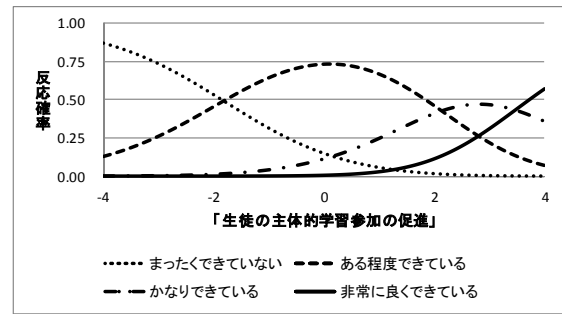


図 24 問 34(7)の項目反応カテゴリ特性曲線（一般化部分採点モデル）

第一に、平均的な自己効力感の教員として、各側面で潜在特性値が 0（偏差値換算で 50）を想定した場合の、三つの側面で共通する結果を述べる。表 5 から、いずれの $\tau_{1,i}$ も高度に有意 ($p < .001$) であり、推定値はいずれも負で、絶対値も大きかった。各図においても、「まったくできていない」と「ある程度できている」についての曲線を比べれば、潜在特性値が 0 の場合の反応確率は前者の方がいずれも低く、ほぼ 0 であった。したがって、日本の前期中等教育段階における平均的な自己効力感の教員は、三つの側面のどの項目でも、最も否定的な「まったくできていない」と「ある程度できている」の回答カテゴリの間では、後者をより選ぶ傾向にあった。

また、いずれの $\tau_{3,i}$ も高度に有意 ($p < .001$) であり、推定値はいずれも正で、絶対値も大きかった。各図においても、「非常に良くできている」と「かなりできている」についての曲線を比べれば、潜在特性値が 0 の場合の反応確率は前者の方がいずれも低く、ほぼ 0 であった。したがって、日本の前期中等教育段階における平均的な自己効力感の教員は、三つの側面のどの項目でも、最も肯定的な「非常に良くできている」と次に肯定的な「かなりできている」の回答カテゴリの間では、前者をより選ぶ傾向にはなかった。

第二に、平均的な自己効力感の教員について、各側面での特徴的な結果を述べる。「学級運営」については、 $\tau_{2,i}$ の絶対値が 0 付近であり、比較的大きい標本サイズであるがいずれも有意ではなかった。図 13 から図 16 においても、「ある程度できている」と「かなりできている」についての曲線を比べると、潜在特性値が 0 付近でいずれも同程度の反応確率であった。したがって、反応率の集計ではいずれも「ある程度できている」が高いものの、本モデルの推定結果では、「学級運営」について日本の前期中等教育段階における平均的な自己効力感の教員は、「学級運営」を測定している項目群において、「ある程度できている」と「かなりできている」の回答カテゴリの間で、分かれた回答をする傾向にあった。

また、「生徒の主体的学習参加の促進」については、いずれの $\tau_{2,i}$ も高度に有意 ($p < .001$) であり、推定値はいずれも正で、絶対値も大きかった。図 21 から図 24 においても、「ある程度できている」と「かなりできている」についての曲線を比べれば、潜在特性値が 0 の場合の反応確率は前者の方がいずれも高かった。したがって、「生徒の主体的学習参加の促進」について日本の前期中等教育段階における平均的な自己効力感の教員は、「生徒の主体的学習参加の促進」を測定している項目群において、「ある程度できている」と「かなりできている」の回答カテゴリの間では、前者をより選ぶ傾向にあった。

なお、「教科指導」については、(11)以外は $\tau_{2,i}$ が正であり、高度に有意 ($p < .001$) であった。図 19 以外の図 17、図 18、図 20 において、潜在特性値が 0 の場合の反応確率は、「ある程度できている」が相対的に高かった。

第三に、自己効力感が相対的に高い教員として、潜在特性値が 1（偏差値換算で 60）を想定した場合の結果を述べる。こうした教員が反応する回答カテゴリの傾向について、「学級運営」と「教科指導」においてはどの項目でも、「かなりできている」と回答する確率が最も高かった。しかしながら、「生徒の主体的学習参加の促進」においては、複数の項目（(1)、(7)）で「かなりできている」ではなく「ある程度できている」への反応確率が最も高かった。(5)では、「ある程度できている」と「かなりできている」への反応確率がいずれも約 49%であった。

第四に、自己効力感が相対的に低い教員として、潜在特性値が-1（偏差値換算で 40）を想定した場合の結果を述べる。こうした教員が反応する回答カテゴリの傾向について、三つの側面のどの項目でも、「ある程度できている」と回答する確率が最も高かった。

5. 考察

本研究では、日本の前期中等教育段階における教員の自己効力感が各側面でどの回答カテゴリに反応する傾向にあるのかについて、自己効力感の個人差との関連で検証することを目的とした。

分析の結果、日本の前期中等教育段階における平均的な自己効力感の教員は、教員の自己効力感に関する三つの側面のどの項目でも、「まったくできていない」と回答する傾向にもなければ、「非常に良くできている」と回答する傾向にもないことが示された。また、特に「学級運営」に関する自己効力感について、日本の前期中等教育段階における平均的な教員は、どの項目でも、「ある程度できている」と「かなりできている」との間で分かれた回答の傾向であることが示された。さらに、特に「生徒の主体的学習参加の促進」に関する自己効力感について、日本の前期中等教育段階における平均的な教員は、どの項目でも、「ある程度できている」までの回答の傾向であることが示された。

このように、特に「生徒の主体的学習参加の促進」について、日本の前期中等教育段階における平均的な自己効力感の教員は、「ある程度できている」までの回答の傾向である一方で、「学級運営」については、「ある程度できている」とより肯定的な「かなりできている」の間で分かれた回答の傾向であることが示唆された。このことは、自己効力感がいずれの側面でも低いという他の参加国との比較から得られた結果の解釈が、国内のデータを用いた分析によって回答カテゴリの反応確率と対比させて得られた結果の解釈とは必ずしも一致しないという点で、TALIS のような国際比較調査の結果を見るに当たっての重要な示唆の一つとなるであろう。

次に、教員の自己効力感が偏差値換算で 60 と相対的に高い場合は、「学級運営」と「教科指導」のどの項目でも、「かなりできている」と回答する確率が高いものの、「生徒の主体的学習参加の促進」については複数の項目で「かなりできている」ではなく「ある程度できている」に回答する確率が高いことが分かった。このことから、「生徒の主体的学習参加の促進」について、平均から 1 標準偏差程度高い水準であっても、「かなりできている」というより肯定的なカテゴリには回答しにくい傾向にある項目があることが示唆された。その一方で、教員の自己効力感が偏差値換算で 40 と相対的に低い場合は、「まったくできていない」ではなく「ある程度できている」に回答する確率が三つの側面のどの項目でも最も高いことが分かった。以上のような回答者の個人差と回答カテゴリの反応確率を関連付けた知見は、回答カテゴリの反応率の集計からでは得られないものである。

本研究で得られた知見の、教員の自己効力感の向上に関する研修への示唆として、以下のことが挙げられよう。「学級運営」では、それを構成する項目群で述べられていることについて、日本の前

期中等教育段階における平均的な自己効力感より1標準偏差程度低い教員は「ある程度できている」と回答し、平均的な自己効力感の教員は、「ある程度できている」と「かなりできている」の間で分かれた回答をし、それより1標準偏差程度高い教員は「かなりできている」と回答する傾向にあると考えられる。このように、「学級運営」については、その個人差が平均 ± 1 標準偏差の間では、回答カテゴリへの反応の違いが見られやすい。その個人差が回答の違いに反映されやすいという点では、「学級運営」の項目群で述べられていることについては教員の個人差に対応した研修は相対的に実施しやすいものと考えられる。他方、「生徒の主体的学習参加の促進」では、それを構成する項目群で述べられていることについて、日本の前期中等教育段階における平均的な自己効力感より1標準偏差程度低い教員のみならず、平均的な自己効力感の教員でも「ある程度できている」に回答する傾向にあり、さらに幾つかの項目群で述べられていることについては、平均より1標準偏差程度高い教員でも「ある程度できている」に回答する傾向にあると考えられる。このように、「生徒の主体的学習参加の促進」については、その個人差が平均 ± 1 標準偏差の間では、回答カテゴリへの反応の違いが見られにくく、相対的に高い自己効力感の教員にとっても、より肯定的な回答をすることが難しい側面である。このことから、「生徒の主体的学習参加の促進」の項目群で述べられていることについては教員の個人差に対応した研修は相対的に実施しにくいものと考えられる。

注釈

- (1) TALIS2013の教員の自己効力感については、「学級運営」、「教科指導」、「生徒の主体的学習参加の促進」の3側面の下位尺度からなっており、それぞれの下位尺度では、異なる4項目ずつが各尺度の測定に用いられた（国立教育政策研究所, 2014）。
- (2) 多次元項目反応モデルの詳細については、Reckase (2009) を参照のこと。
- (3) 確認的な多次元段階反応モデルについて、正規累積モデルで表し、WLSMV と呼ばれる異なる推定方法を用いた場合、Mplusにおいて適合度指標を出力することができる。本研究では、Asparouhov and Muthén (2010) を参考に、制約のない2レベルモデルに基づいた多重代入法を用いて欠測値が代入されたデータセットを五つ作成し、それぞれのデータセットに対して当該のモデルを当てはめ、平均的な適合度指標を得た。その結果、CFI = 0.980、TLI = 0.974、RMSEA = 0.074、WRMR = 2.717 であり、WRMR 以外は低い適合度であるとは言えなかった。
- (4) 多次元項目反応モデルにおいて、例えば二つの潜在特性がある項目に影響を与えることを仮定し、この二つの潜在特性の大小によって反応確率がどのように変化するかを視覚的に表現したい場合、3次元の図を用いて項目反応曲面を描く。本研究では、各項目に影響を与える潜在特性はいずれか一つだけであるという確認的なモデルを仮定しているため、解釈が容易である、対応する一つの潜在特性との関係を平面上に描く項目反応カテゴリ特性曲線を用いている。
- (5) (4)と同様。

引用・参考文献

Asparouhov, T. & Muthén, B. (2010). *Multiple imputation with Mplus (Version 2)*. Retrieved November 27, 2015, from <http://www.statmodel.com/download/Imputations7.pdf>.

Asparouhov, T. & Muthén, B. (2015). *IRT in Mplus*. Retrieved November 26, 2015, from <http://www.statmodel.com/download/MplusIRT.pdf>.

国立教育政策研究所（編）（2014）. 教員環境の国際比較：OECD 国際教員指導環境調査（TALIS）2013年調査結果報

告書 明石書店.

- Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement*, **16**, 159-176.
- Muraki, E., & Carlson, J. E. (1995). Full-information factor analysis for polytomous item responses. *Applied Psychological Measurement*, **19**, 73-90.
- Muthén, L. K., & Muthén, B. O. (1998–2015). *Mplus user's guide. Seventh edition*. Los Angeles, CA: Muthén & Muthén.
- OECD (2014a). *TALIS 2013 results: An international perspective on teaching and learning*. OECD Publishing.
- OECD (2014b). *TALIS 2013 technical report*. Retrieved October 22, 2014, from <http://www.oecd.org/edu/school/TALIS-technical-report-2013.pdf>.
- Reckase, M. D. (2009). *Multidimensional item response theory*. New York: Springer.
- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, **63**, 581-592.
- Samejima, F. (1969). *Estimation of latent ability using a response pattern of graded scores* (Psychometric Monograph No. 17). Richmond, VA: Psychometric Society.
- Yao, L., & Schwarz, R. D. (2006). A multidimensional partial credit model with associated item and test statistics: An application to mixed-format tests. *Applied Psychological Measurement*, **30**, 469-492.

付記・謝辞

※本研究では以下の URL にある TALIS 2013 の Complete Database を用いた。

http://stats.oecd.org/Index.aspx?datasetcode=talis_2013%20

※本研究は、日本教育心理学会第 57 回総会にてポスター発表したものに分析を追加し、加筆修正したものである。

※本研究の一部は JSPS 科研費 JP26350219 の助成を受けたものである。

(受理日：平成 28 年 3 月 31 日)