

RESEARCH

Open Access



A statistical analysis of the characteristics of the intended curriculum for Japanese primary science and its relationship to the attained curriculum

Kenji Matsubara^{1*}, Yasuhito Hagiwara¹ and Yuji Saruta²

*Correspondence:
kmatsuba@nier.go.jp

¹ National Institute
for Educational Policy
Research, 3-2-2,
Kasumigaseki, Chiyoda-ku,
Tokyo, Japan
Full list of author information
is available at the end of the
article

Abstract

This study statistically investigates the characteristics of the intended curriculum for Japanese primary science, focusing on the learning content. The study used the TIMSS 2011 Grade 4 Curriculum Questionnaire data as a major source for the learning content prescribed at the national level. Confirmatory factor analysis was used to determine the extent to which a topic area was covered, as compared to the average among the 59 TIMSS 2011 participating countries. The study revealed that the topic areas of “Human Health” and “Changes in Environments,” both in the life science domain, showed statistically less coverage in the Japanese primary science curriculum when compared to the international average. Furthermore, in discussion, the study relates the characteristics found in the intended curriculum to those in the attained curriculum, examining the percent correct statistics for relevant items from the science assessment. Based on these findings, the study proposes two recommendations for revision of the Japanese primary science curriculum.

Keywords: Curriculum revision, TIMSS, Curriculum Questionnaire, Japanese primary science curriculum, CT-CM model, Bayesian analysis, Plausible values

Background

Japan’s national curricula, the “Courses of Study,” have been revised roughly once every 10 years by the Ministry of Education, Culture, Sports, Science and Technology (MEXT) since first being implemented in 1947. The current Course of Study for Elementary Schools was announced in March 2008, and came fully into force in April 2011 with the objectives of revision including keeping up with societal change over the years and meeting the needs of learners on an international level. The Course of Study for Elementary Schools sets out the overall goals, objectives, and content for each subject and learning area.

Textbooks in Japan must be authorized by MEXT, and their drafts are always checked against the content described in the Courses of Study. Since these textbooks are extensively used by primary school teachers throughout Japan, this system functions to maintain the levels of learning content expected for both teaching and learning at schools.

Thus, selecting the appropriate learning content at the right grade level is clearly of great importance when it comes to revising the national curricula, particularly in the case of Japan's Course of Study for Elementary Schools. Multiple perspectives should be considered when judging the appropriateness of learning content for a particular grade; one such perspective that can be particularly useful is international comparison of learning content.

One of the models used to frame curriculum studies is the well-known curriculum model of the TIMSS, which comprises three aspects: the intended, the implemented, and the attained curriculum (Mullis et al. 2009). In this study, Japan's primary science curriculum is regarded as the model's intended curriculum since it is part of the national curriculum. The primary science textbooks, used extensively by the teachers, are regarded as the implemented curriculum whereas the pupils' performance or achievement on standardized tests is regarded as the achieved curriculum. Using this curriculum model and the available data, the learning content described in the national (intended) curriculum can be compared to an international standard. It should be noted that while the context of an individual country's particular learning content can be dealt with in terms of both the implemented and attained curriculum, this study will discuss only the attained curriculum.

The TIMSS 2011 Curriculum Questionnaire provides important information necessary for an international comparison of primary science curricula at the intended curriculum level. It offers up-to-date information on which topic areas (units of learning content) are taught in science curricula by grade in all the TIMSS 2011 participating countries. The TIMSS 2003 Curriculum Questionnaire, which surveyed National Research Coordinators about the grades in which particular science topic areas were expected to be taught, included 53 topic areas and reported primarily on general international trends of change in science curricula (Martin et al. 2004). Using data from past cycles and the TIMSS 2003 international data, Saruta (2007) summarized trends in Japan's science curriculum since the 1960s, and pointed out several characteristics of the Japanese national science curriculum on a topic-area basis. The results of this study contributed to development and revision of Japan's current science curriculum. Saruta's study counted the number of countries in which a topic was taught based on both topic area and grade, and used this as a basis for comparison. Although this produced some important information, a stronger means of international comparison is required. This study, with the help of statistical analysis, attempts to determine the extent to which topic areas are covered taking into account international perspectives.

Purpose

This study investigates the characteristics of the Japanese primary science curriculum through an international comparison of learning content. It is expected to contribute to the upcoming revision of the national curriculum. Life science was chosen as the focus domain. The study addresses three research questions: (1) Which learning content areas in life science are less thoroughly covered by Japan's primary-level science curriculum when compared to other countries' primary-level science curricula? (2) How do Japanese pupils at the primary level perform in these less-covered learning content areas? (3) What recommendations can be made for future revision of the Japanese primary science

curriculum, considering the content covered by the national curriculum and the performance of the pupils?

Methods

Data sources

For internationally comparable information about learning content at the intended curriculum level, this study used the results of the TIMSS 2011 Grade 4 Curriculum Questionnaire (SC408A), which provides information on 59 participating countries and regions. The questionnaire surveyed participants about the grades (from preprimary through upper secondary school) in which certain topics in the life science domain are intended to be taught. Results involving preprimary (PP) and G7 to G12 levels were excluded since this study is concerned with revision of the primary-level education curriculum. Table 1 shows the six topics and their descriptions as they appear in SC408A.

For internationally comparable information about less-thoroughly covered learning content in Japan's primary science curriculum at the attained curriculum level, this study used the percent correct statistics for items from the TIMSS 2011 science assessment almanacs data, along with the item information and TCMA (Test Curriculum Matching Analysis) data (Martin et al. 2012).

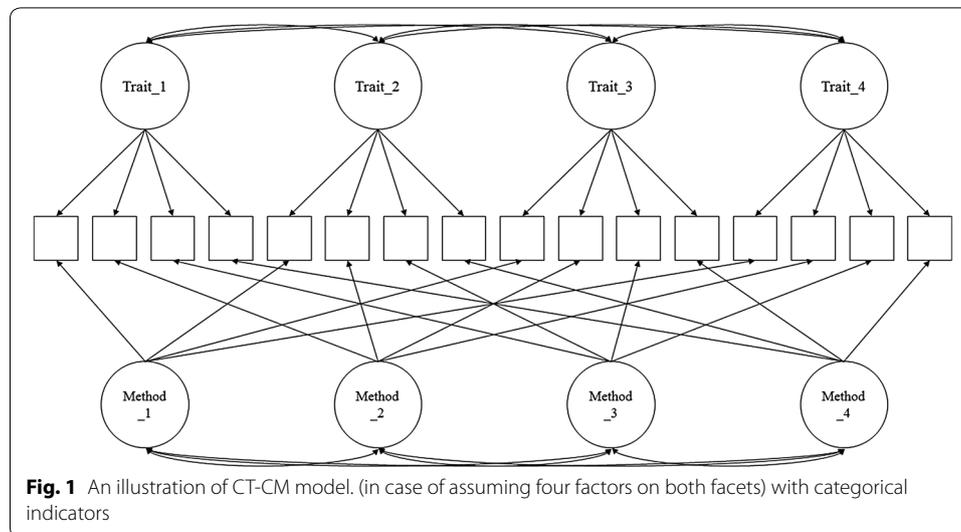
Statistical model used in this study

To examine which learning content areas in life sciences are less thoroughly covered by Japan's primary science curriculum in comparison to other countries' science curricula, we analyzed the data on the content domain of life science at the primary school level using an additive model based on confirmatory factor analysis. Specifically, as in an analysis of the multitrait-multimethod matrix (Campbell and Fiske 1959), we use the CT-CM (correlated trait-correlated method) model (e.g., Widaman 1985) with categorical (binary) indicators, treating topics as trait factors and grades as method factors (see Fig. 1).

To investigate which of the six topics in life science are (more or) less covered in a country's science curriculum, we need to measure the degree to which the country covers each topic, after controlling for the effect of the curriculum coverage of life science in each grade. As Fig. 1 shows, the CT-CM model assumes that each item measures one of the factors for traits (topics in this study) and one of the factors for methods (grades in this study) independently. From this point, applying this model to the TIMSS

Table 1 Topic areas described in the Curriculum Questionnaire (Source: International Version of the TIMSS 2011 Background and Curriculum Questionnaires Supplement 1, pp. 155–156)

SC408AA	a) Major body structures and their functions in humans and other organisms (plants and animals)
SC408AB	b) Life cycles and reproduction in plants and animals
SC408AC	c) Physical features, behavior, and survival of organisms living in different environments
SC408AD	d) Relationships in a given community (e.g., simple food chains, predator prey relationships)
SC408AE	e) Changes in environments (effects of human activity, pollution and its prevention)
SC408AF	f) Human health (e.g., transmission/prevention of communicable diseases, signs of health/illness, diet, exercise)



Curriculum Questionnaire data is thought to be appropriate in light of the purpose of this study.

However, the CT-CM model is known to have identification issues and to result in Heywood cases (Kenny and Kashy 1992; Marsh 1989).¹ To deal with these issues and the small sample size of the TIMSS Curriculum Questionnaire data ($N = 59$), we used a Bayesian estimation. However, a Bayesian estimation of the categorical factor model with binary indicators and with a small sample size is known to show prior assumption dependence (Asparouhov and Muthén 2010b). Therefore, we conducted simulation studies to vary the conditions of prior distributions before real-data analysis. For the analyses of this study, we used Mplus version 7.3 (Muthén and Muthén 1998–2012).

Simulation studies

We assume that 60 countries were asked whether six topics were covered in six of the country's respective grade levels. Therefore, there are 36 observed dichotomous variables per country.

In the CT-CM model, we set all factor loadings to 1, and all factor correlations to 0.6 as true values. We set the true values of threshold parameters as follows: 1.5 for grade 1, 0.9 for grade 2, 0.3 for grade 3, -0.3 for grade 4, -0.9 for grade 5, and -1.5 for grade 6 on all topics. The means and variances of the factors are fixed at 0 and 1, respectively.

We set four conditions on prior distribution of factor loadings and thresholds. Specifically, while Mplus default priors in the categorical factor model are $N(0,5)$, we additionally set $N(0,1)$, $N(0,2.25)$, and $N(0,4)$ conditions, because the simulation study of the categorical factor model for small sample sizes suggests that small variance priors such as $N(0,1)$ may be more preferable (Asparouhov and Muthén 2010b). The prior distribution of the factor correlation matrix on topics was $IW(\mathbf{I}, p + 1)$ (p was the number of

¹ Eid (2000) proposed the so-called correlated trait-correlated method minus one [CT-C ($M - 1$)] model, which always identifies a solution. However, it is known that the solution depends on which method factor is excluded from the model as the reference. As it is not appropriate to choose one particular grade as the reference method, we did not use the CT-C($M - 1$) model.

factors, six in this study) and the prior distribution of grade was also $IW(\mathbf{I}, p + 1)$, the Mplus default priors in the categorical factor analysis.

The number of MCMC (Markov chain Monte Carlo) chains for each parameter was two, and the minimum number of iterations for each chain was 100,000. By inputting THIN = 2 in Mplus, we used every second iteration and discarded the first half of each chain (Muthén and Muthén 1998–2012). The convergence was judged using a potential scale reduction (PSR) (Gelman and Rubin 1992) of less than 1.1. Two hundred simulations were run under each of the four conditions. We used medians as point estimates of the parameters.

Real-data analysis

We used the four conditions shown in the previous section. The number of MCMC chains for each parameter was five, and the minimum number of iterations for each chain was 100,000. By inputting THIN = 5 in Mplus, we used every fifth iteration and discarded the first half of each chain. The convergence criterion was the same as in the simulation study.

Results

Results of the simulation studies

All 800 runs were successful computations according to the output files. We evaluated the results in terms of parameter bias and 95 % coverage. The latter is the proportion of replications for which the 95 % confidence interval contains the population parameter value (Muthén and Muthén 1998–2012). Table 2 shows the minimum, average, and maximum bias, and Table 3 shows the 95 % coverage for each kind of parameter (factor loadings, factor correlations, and thresholds).

Underestimation was found on factor loadings in the $N(0,1)$ condition, and overestimation was found on factor loadings and thresholds in the $N(0,4)$ and $N(0,5)$ conditions. Underestimation on factor correlations was found in all conditions, to a higher degree as prior variance was bigger. Conversely, the coverages were close to 95 % (over 90 %) in

Table 2 Parameter bias in the simulation studies

	$N(0,1)$	$N(0,2.25)$	$N(0,4)$	$N(0,5)$
Factor loadings				
Min	−0.343	−0.106	0.086	0.166
Mean	−0.200	0.042	0.232	0.314
Max	−0.106	0.155	0.354	0.445
Thresholds				
Min	−0.213	−0.102	−0.336	−0.440
Mean	0.003	0.004	0.006	0.006
Max	0.221	0.103	0.347	0.459
Correlations between factors				
Min	−0.091	−0.134	−0.155	−0.159
Mean	−0.071	−0.114	−0.130	−0.135
Max	−0.043	−0.094	−0.108	−0.111

Table 3 95 % coverage in the simulation studies

	$N(0,1)$	$N(0,2.25)$	$N(0,4)$	$N(0,5)$
Factor loadings				
Min	96.0	93.0	90.5	88.5
Mean	98.2	97.3	95.2	94.0
Max	100.0	100.0	99.5	98.5
Thresholds				
Min	94.5	93.5	93.0	93.0
Mean	96.5	97.5	96.6	95.9
Max	99.0	100.0	99.5	99.0
Correlations between factors				
Min	91.0	91.5	90.5	90.5
Mean	93.9	95.4	94.6	94.1
Max	96.0	99.5	99.0	98.5

almost all conditions, except for some factor loadings in the $N(0,5)$ condition for which coverages below 90 % were found.

These results suggest that the point estimate is inappropriate even when bigger or smaller variances are used for prior distributions of factor loadings and thresholds. Instead, interval estimation may be more preferable and appropriate, as its results seem robust over a range of prior distributions. Therefore, in the real-data analysis in the next section, we will show the results from all four conditions in the simulation study, but will primarily report the results of the credible interval estimation.

Results of the real-data analysis

After 100,000 iterations, PSRs were less than 1.01 for all conditions and the posterior predictive p value was 0.348 for the $N(0,1)$ condition, 0.374 for the $N(0,2.25)$ condition, 0.381 for the $N(0,4)$ condition, and 0.382 for the $N(0,5)$ condition. Additionally, R-squared values did not exceed 1 for any of the 36 indicators or factor correlations in all conditions. From these results, we judged that all model estimations terminated normally and that goodness-of-fit was high to some extent, particularly for the latter three conditions.² The 95 % credible intervals of each parameter for each condition are shown in Table 4. Results showed that none of the factor loadings for the topic factors had negative intervals, and that more than half of these had positive intervals in all conditions.

We obtained each country's posterior distributions of the factors for each topic area and for each grade, drawing 10,000 plausible values per factor and per country. Asparouhov and Muthén (2010a, p. 2) said, "...the individual level latent variable is of interest. One such example is the case when we want to construct a factor scores estimate for a latent variable and a standard error for that factor score estimate. To be able to do this it is necessary to use many imputed values." They also "conclude that for small sample size the Bayes factor score estimates and standard errors are more reliable than those

² In the frequentist approach using a WLSMV (weighted least square parameter estimates using a diagonal weight matrix) estimator and Delta parameterization in Mplus, the model estimation terminated normally, but a Heywood case occurred (R-squared values of some indicators exceeded 1).

Table 4 Credible intervals of each parameter

	<i>N</i> (0,1)		<i>N</i> (0,2.25)		<i>N</i> (0,4)		<i>N</i> (0,5)	
	95 % CI		95 % CI		95 % CI		95 % CI	
	Lower 2.5 %	Upper 2.5 %	Lower 2.5 %	Upper 2.5 %	Lower 2.5 %	Upper 2.5 %	Lower 2.5 %	Upper 2.5 %
Factor loadings								
FA_BY								
SC408AAB	0.898	3.015	1.199	4.503	1.349	5.876	1.436	6.563
SC408AAC	0.947	3.095	1.252	4.536	1.420	5.854	1.528	6.437
SC408AAD	0.437	2.126	0.569	3.086	0.645	4.025	0.688	4.497
SC408AAE	-0.113	1.283	-0.054	1.607	-0.049	1.811	-0.042	1.891
SC408AAF	-0.904	0.613	-1.048	0.677	-1.145	0.709	-1.180	0.726
SC408AAG	-0.956	0.427	-1.132	0.519	-1.273	0.645	-1.354	0.675
FB_BY								
SC408ABB	0.531	2.409	0.329	3.391	0.022	4.380	0.010	4.857
SC408ABC	0.664	2.754	0.347	3.835	-0.092	4.816	-0.178	5.277
SC408ABD	0.298	1.921	0.242	2.693	0.057	3.467	0.032	3.869
SC408ABE	-0.227	1.158	-0.306	1.449	-0.391	1.701	-0.423	1.836
SC408ABF	-0.794	0.709	-0.919	0.831	-0.978	0.964	-1.003	1.052
SC408ABG	-0.857	1.097	-1.180	1.210	-1.572	1.326	-1.727	1.384
FC_BY								
SC408ACB	0.764	2.873	0.958	4.220	1.064	5.467	1.123	6.042
SC408ACC	0.837	3.053	1.067	4.480	1.181	5.825	1.246	6.443
SC408ACD	0.430	1.971	0.337	2.381	0.163	2.672	0.135	2.801
SC408ACE	0.170	1.738	-0.028	2.175	-0.252	2.507	-0.296	2.640
SC408ACF	-0.601	0.797	-0.836	0.788	-1.134	0.764	-1.193	0.761
SC408ACG	-1.426	0.371	-1.842	0.305	-2.207	0.247	-2.351	0.232
FD_BY								
SC408ADB	0.395	1.970	0.462	2.798	0.459	3.680	0.496	4.093
SC408ADC	0.918	3.124	0.836	4.350	0.537	5.442	0.568	5.999
SC408ADD	0.389	1.882	0.345	2.417	0.243	3.004	0.234	3.295
SC408ADE	0.183	1.968	0.139	2.749	0.030	3.466	0.036	3.796
SC408ADF	-0.628	0.906	-0.852	1.017	-1.156	1.097	-1.283	1.141
SC408ADG	-0.922	0.709	-1.170	0.821	-1.380	0.907	-1.468	0.937
FE_BY								
SC408AEB	0.894	2.826	1.169	4.084	1.348	5.231	1.433	5.794
SC408AEC	0.794	2.668	1.061	3.916	1.253	5.146	1.347	5.706
SC408AED	0.927	3.120	1.051	4.415	1.091	5.608	1.129	6.160
SC408AEE	0.024	1.509	0.085	2.110	0.132	2.682	0.159	2.996
SC408AEF	-0.494	0.803	-0.532	0.982	-0.581	1.104	-0.602	1.164
SC408AEG	-0.741	0.511	-0.741	0.608	-0.747	0.681	-0.746	0.714
FF_BY								
SC408AFB	1.177	3.423	1.399	4.854	1.483	6.154	1.551	6.787
SC408AFC	0.795	2.670	0.913	3.701	0.918	4.682	0.928	5.166
SC408AFD	0.986	3.068	1.230	4.382	1.373	5.610	1.441	6.231
SC408AFE	0.490	2.093	0.670	3.019	0.798	4.054	0.843	4.530
SC408AFF	-0.084	1.119	-0.006	1.366	0.041	1.541	0.055	1.598
SC408AFG	-0.138	0.897	-0.065	1.128	-0.020	1.387	-0.006	1.479
F_BY								
SC408AAB	-0.190	1.995	-0.449	2.910	-0.632	3.882	-0.726	4.288
SC408ABB	-0.316	1.287	-0.427	2.279	-0.507	3.454	-0.564	3.911

Table 4 continued

	<i>N</i> (0,1)		<i>N</i> (0,2.25)		<i>N</i> (0,4)		<i>N</i> (0,5)	
	95 % CI		95 % CI		95 % CI		95 % CI	
	Lower 2.5 %	Upper 2.5 %	Lower 2.5 %	Upper 2.5 %	Lower 2.5 %	Upper 2.5 %	Lower 2.5 %	Upper 2.5 %
SC408ACB	-0.690	1.127	-1.017	1.975	-1.248	2.934	-1.363	3.316
SC408ADB	-0.257	1.233	-0.467	1.791	-0.623	2.425	-0.731	2.654
SC408AEB	-0.243	1.521	-0.346	2.415	-0.398	3.441	-0.430	3.876
SC408AFB	-0.222	2.289	-0.442	3.410	-0.612	4.648	-0.679	5.191
F_C BY								
SC408AAC	-0.128	2.189	-0.307	3.180	-0.375	4.253	-0.401	4.700
SC408ABC	-0.084	1.835	-0.146	2.979	-0.136	4.224	-0.155	4.723
SC408ACC	-0.701	1.226	-1.058	2.121	-1.370	3.114	-1.477	3.494
SC408ADC	-0.688	1.171	-0.937	2.038	-1.194	2.955	-1.313	3.297
SC408AEC	-0.263	1.507	-0.392	2.354	-0.455	3.254	-0.512	3.661
SC408AFC	-0.208	2.214	-0.382	3.205	-0.459	4.222	-0.494	4.643
F_D BY								
SC408AAD	0.141	2.227	0.048	3.203	0.143	4.116	0.169	4.531
SC408ABD	0.011	1.618	-0.038	2.602	0.048	3.764	0.060	4.245
SC408ACD	-0.291	1.116	-0.353	1.632	-0.384	2.168	-0.396	2.345
SC408ADD	-0.222	1.096	-0.267	1.654	-0.264	2.307	-0.271	2.622
SC408AED	-0.630	1.102	-0.829	1.884	-0.906	2.792	-0.946	3.217
SC408AFD	0.014	2.152	-0.083	2.958	-0.114	3.810	-0.131	4.168
F_E BY								
SC408AAE	0.505	2.238	0.541	2.755	0.572	3.160	0.579	3.318
SC408ABE	0.522	2.115	0.658	2.842	0.728	3.480	0.746	3.727
SC408ACE	0.425	2.038	0.580	2.813	0.670	3.434	0.691	3.696
SC408ADE	0.777	2.764	1.086	4.073	1.290	5.287	1.366	5.862
SC408AEE	0.668	2.544	0.957	3.707	1.147	4.817	1.227	5.330
SC408AFE	0.191	1.677	0.181	2.170	0.177	2.713	0.171	2.959
F_F BY								
SC408AAF	0.733	2.616	0.790	3.429	0.805	4.112	0.803	4.398
SC408ABF	0.818	2.578	0.963	3.471	1.033	4.237	1.052	4.577
SC408ACF	0.623	2.064	0.737	2.642	0.804	3.118	0.825	3.311
SC408ADF	0.800	2.520	0.953	3.489	1.046	4.470	1.076	4.961
SC408AEF	0.613	2.147	0.717	2.810	0.783	3.417	0.803	3.698
SC408AFF	0.560	1.984	0.589	2.387	0.600	2.683	0.600	2.782
F_G BY								
SC408AAG	0.744	2.545	0.901	3.496	0.990	4.381	1.028	4.780
SC408ABG	1.175	3.338	1.440	4.668	1.638	5.903	1.729	6.510
SC408ACG	1.189	3.286	1.466	4.563	1.641	5.711	1.709	6.265
SC408ADG	0.952	2.756	1.138	3.693	1.252	4.542	1.290	4.909
SC408AEG	0.478	1.769	0.532	2.047	0.566	2.233	0.576	2.293
SC408AFG	0.343	1.419	0.364	1.625	0.377	1.806	0.382	1.881
Correlations between factors								
FB_ WITH								
FA_	-0.070	0.892	-0.459	0.836	-0.679	0.807	-0.673	0.799
FC_ WITH								
FA_	0.323	0.932	0.073	0.898	-0.101	0.878	-0.108	0.872
FB_	0.471	0.956	0.110	0.935	-0.195	0.919	-0.224	0.914
FD_ WITH								

Table 4 continued

	<i>N</i> (0,1)		<i>N</i> (0,2.25)		<i>N</i> (0,4)		<i>N</i> (0,5)	
	95 % CI		95 % CI		95 % CI		95 % CI	
	Lower 2.5 %	Upper 2.5 %	Lower 2.5 %	Upper 2.5 %	Lower 2.5 %	Upper 2.5 %	Lower 2.5 %	Upper 2.5 %
FA_	0.133	0.903	-0.142	0.850	-0.319	0.822	-0.312	0.816
FB_	0.555	0.963	0.210	0.950	-0.129	0.940	-0.147	0.937
FC_	0.569	0.960	0.328	0.944	0.119	0.931	0.102	0.927
FE_WITH								
FA_	0.489	0.955	0.297	0.926	0.134	0.898	0.101	0.889
FB_	-0.230	0.804	-0.511	0.724	-0.693	0.686	-0.693	0.676
FC_	0.114	0.875	-0.061	0.823	-0.196	0.801	-0.211	0.795
FD_	0.013	0.825	-0.170	0.757	-0.337	0.730	-0.356	0.722
FF_WITH								
FA_	0.557	0.958	0.466	0.940	0.394	0.929	0.382	0.927
FB_	-0.262	0.804	-0.549	0.732	-0.719	0.699	-0.711	0.689
FC_	0.044	0.873	-0.151	0.811	-0.276	0.782	-0.286	0.774
FD_	-0.021	0.826	-0.189	0.766	-0.322	0.744	-0.326	0.739
FE_	0.684	0.968	0.554	0.955	0.451	0.943	0.422	0.938
F_C WITH								
F_B	-0.385	0.975	-0.497	0.966	-0.469	0.963	-0.479	0.962
F_D WITH								
F_B	-0.411	0.974	-0.481	0.960	-0.446	0.952	-0.442	0.949
F_C	-0.222	0.974	-0.324	0.959	-0.268	0.952	-0.274	0.950
F_E WITH								
F_B	-0.543	0.967	-0.589	0.939	-0.532	0.923	-0.537	0.918
F_C	-0.360	0.969	-0.397	0.948	-0.348	0.936	-0.343	0.932
F_D	0.202	0.968	0.060	0.944	0.098	0.931	0.096	0.925
F_F WITH								
F_B	-0.563	0.973	-0.591	0.957	-0.499	0.937	-0.502	0.928
F_C	-0.396	0.972	-0.396	0.952	-0.325	0.928	-0.325	0.917
F_D	0.167	0.972	-0.001	0.953	0.070	0.931	0.061	0.921
F_E	0.445	0.963	0.384	0.916	0.368	0.894	0.359	0.887
F_G WITH								
F_B	-0.565	0.974	-0.596	0.957	-0.498	0.938	-0.506	0.928
F_C	-0.379	0.974	-0.387	0.956	-0.313	0.936	-0.316	0.926
F_D	0.157	0.973	-0.068	0.955	0.014	0.934	0.011	0.924
F_E	0.482	0.964	0.420	0.919	0.403	0.894	0.395	0.886
F_F	0.792	0.974	0.751	0.962	0.731	0.955	0.724	0.953
Thresholds								
SC408AAB	-0.822	0.540	-1.109	0.761	-1.383	0.969	-1.517	1.060
SC408AAC	-1.098	0.330	-1.515	0.445	-1.928	0.552	-2.117	0.598
SC408AAD	-1.260	-0.001	-1.713	-0.019	-2.165	-0.033	-2.381	-0.038
SC408AAE	-1.720	-0.449	-2.037	-0.500	-2.296	-0.531	-2.394	-0.542
SC408AAF	-1.521	-0.262	-1.851	-0.298	-2.126	-0.318	-2.237	-0.323
SC408AAG	-1.088	0.041	-1.392	0.015	-1.654	0.001	-1.772	0.000
SC408ABB	0.473	1.791	0.549	2.390	0.607	3.048	0.629	3.392
SC408ABC	-0.565	0.683	-0.760	0.857	-0.945	1.029	-1.038	1.108
SC408ABD	-1.014	0.072	-1.347	0.074	-1.706	0.087	-1.875	0.093
SC408ABE	-1.487	-0.322	-1.857	-0.374	-2.173	-0.410	-2.301	-0.422
SC408ABF	-0.972	0.167	-1.234	0.162	-1.464	0.151	-1.563	0.150

Table 4 continued

	<i>N</i> (0,1)		<i>N</i> (0,2.25)		<i>N</i> (0,4)		<i>N</i> (0,5)	
	95 % CI		95 % CI		95 % CI		95 % CI	
	Lower 2.5 %	Upper 2.5 %	Lower 2.5 %	Upper 2.5 %	Lower 2.5 %	Upper 2.5 %	Lower 2.5 %	Upper 2.5 %
SC408ABG	-0.828	0.536	-1.129	0.630	-1.422	0.714	-1.570	0.755
SC408ACB	-0.232	0.969	-0.314	1.336	-0.398	1.704	-0.430	1.880
SC408ACC	-0.447	0.807	-0.608	1.118	-0.760	1.436	-0.826	1.574
SC408ACD	-0.813	0.176	-0.906	0.201	-0.987	0.221	-1.026	0.227
SC408ACE	-1.235	-0.082	-1.508	-0.088	-1.742	-0.096	-1.843	-0.099
SC408ACF	-0.481	0.502	-0.585	0.539	-0.688	0.563	-0.727	0.572
SC408ACG	-0.490	0.871	-0.735	1.014	-0.981	1.111	-1.094	1.163
SC408ADB	0.875	2.310	0.994	3.006	1.079	3.776	1.113	4.130
SC408ADC	0.298	1.748	0.357	2.306	0.391	2.829	0.404	3.082
SC408ADD	-0.110	0.839	-0.127	0.979	-0.141	1.147	-0.147	1.234
SC408ADE	-1.106	0.219	-1.553	0.243	-2.004	0.247	-2.221	0.249
SC408ADF	-0.415	0.693	-0.572	0.781	-0.735	0.863	-0.811	0.900
SC408ADG	-0.375	0.816	-0.516	0.935	-0.651	1.021	-0.706	1.051
SC408AEB	0.795	2.467	0.989	3.363	1.154	4.227	1.220	4.615
SC408AEC	0.453	1.899	0.545	2.588	0.626	3.294	0.655	3.624
SC408AED	-0.711	0.582	-0.920	0.761	-1.124	0.944	-1.232	1.031
SC408AEE	-1.718	-0.359	-2.398	-0.477	-3.053	-0.576	-3.367	-0.620
SC408AEF	-0.832	0.189	-1.039	0.178	-1.225	0.171	-1.309	0.169
SC408AEG	-0.654	0.241	-0.744	0.229	-0.811	0.222	-0.834	0.219
SC408AFB	-0.331	1.239	-0.456	1.650	-0.597	2.029	-0.674	2.195
SC408AFC	-0.491	0.834	-0.648	1.063	-0.808	1.265	-0.879	1.351
SC408AFD	-1.206	0.250	-1.612	0.311	-1.997	0.364	-2.185	0.387
SC408AFE	-1.287	-0.100	-1.640	-0.128	-2.006	-0.139	-2.178	-0.140
SC408AFF	-0.643	0.360	-0.789	0.352	-0.899	0.343	-0.938	0.341
SC408AFG	-0.340	0.503	-0.389	0.520	-0.430	0.540	-0.445	0.547

FA_: major body structures and their functions in humans and other organisms

FB_: life cycles and reproduction in plants and animals

FC_: physical features, behavior, and survival of organisms living in different environments

FD_: relationships in a given community

FE_: changes in environments

FF_: human health

F_B: grade 1, F_C: grade 2, F_D: grade 3, F_E: grade 4, F_F: grade 5, F_G: grade 6

obtained by the ML estimator” (p. 5). Hence, we drew each country’s imputed plausible values for latent variables and computed their 95 % intervals.

To confirm the accuracy of plausible values in each condition, we show the minimum, average, and maximum means and variances of 12 factors in Table 5. These results tell us that the estimates of the variances in the $N(0,1)$ condition were obviously overestimated. Variance estimates in this condition were thus less accurate than in the other conditions.

Japan’s 95 % intervals for the plausible values of the factors are shown in Table 6. The result showed negative intervals in all four conditions for Japan’s factors on the topics of “Human Health” and “Changes in Environments.” This means that the coverages of the Japanese primary science curriculum on these two topics are below the international averages, after controlling for the effect of the curriculum coverage of life science in each grade. In the $N(0,1)$

Table 5 Means and variances of plausible values for the factors

	<i>N</i> (0,1)	<i>N</i> (0,2.25)	<i>N</i> (0,4)	<i>N</i> (0,5)
Factor means (we set the values as 0 in the model)				
Min	-0.070	-0.052	-0.043	-0.041
Mean	0.017	0.010	0.006	0.004
Max	0.108	0.079	0.061	0.054
Factor variances (we set the values as 1 in the model)				
Min	1.072	0.998	0.962	0.951
Mean	1.127	1.036	0.996	0.984
Max	1.180	1.072	1.023	1.008

These statistics are based on the average results over 10,000 data sets

Table 6 95 % intervals of Japan's plausible values for each factor

	<i>N</i> (0,1)		<i>N</i> (0,2.25)		<i>N</i> (0,4)		<i>N</i> (0,5)	
	95 % CI		95 % CI		95 % CI		95 % CI	
	Lower 2.5 %	Upper 2.5 %	Lower 2.5 %	Upper 2.5 %	Lower 2.5 %	Upper 2.5 %	Lower 2.5 %	Upper 2.5 %
FA_	-2.325	-0.092	-2.096	0.132	-1.954	0.387	-1.908	0.452
FB_	-2.065	0.359	-1.895	0.718	-1.803	1.091	-1.768	1.191
FC_	-2.417	-0.226	-2.328	0.010	-2.294	0.228	-2.286	0.297
FD_	-2.416	-0.006	-2.371	0.212	-2.348	0.449	-2.346	0.507
FE_	-2.621	-0.165	-2.636	-0.111	-2.643	-0.053	-2.637	-0.039
FF_	-2.559	-0.291	-2.474	-0.248	-2.402	-0.209	-2.385	-0.189
F_B	-1.074	1.002	-1.380	1.123	-1.528	1.173	-1.588	1.195
F_C	-1.062	0.963	-1.382	1.003	-1.495	1.017	-1.540	1.009
F_D	-0.810	1.288	-0.940	1.573	-0.993	1.627	-1.009	1.670
F_E	-0.999	0.786	-1.028	0.674	-1.055	0.606	-1.057	0.589
F_F	-0.777	0.692	-0.807	0.577	-0.820	0.491	-0.823	0.463
F_G	-0.482	0.999	-0.411	0.938	-0.375	0.910	-0.361	0.898

condition, all topic factors except for “Life Cycles and Reproduction in Plants and Animals” also showed negative intervals for Japan. However, this condition does not seem to be relatively well-fitted and it does not provide accurate estimates for factor variances. Therefore, we interpret this to mean that the topic areas of “Changes in Environments” and “Human Health” were relatively less thoroughly covered in the Japanese primary science curriculum compared to the average of the curricula of the 59 TIMSS 2011 participating countries.

From the viewpoint of grades, the results showed the intervals including zero for Japan's factors on grades in all the four conditions. Therefore, we interpret this to mean that coverage of the curriculum in Japanese primary science at each grade level does not differ much compared to the average of coverage in the respective grades of the 59 TIMSS 2011 participating countries.

Discussion

The results of the statistical analysis revealed that the topic areas of “Human Health” and “Changes in Environments” were relatively less covered in the Japanese primary science curriculum. In discussion, the study related the characteristics found in the

learning content in the intended curriculum to the corresponding learning content in the attained curriculum. To compare the achievement of Japanese primary pupils in the less thoroughly covered learning content, we examined the percent correct statistics for each of the released TIMSS 2011 science items. The framework for these science items is described in the TIMSS 2011 Science Framework as follows, showing the content domain of life science and an overview of the topic areas to be covered:

“Life science includes understandings of the characteristics and life processes of living things, the relationships between them, and their interaction with the environment. The topic areas for life science are as follows:

- Characteristics and life processes of living things
- Life cycles, reproduction, and heredity
- Interaction with the environment
- Ecosystems
- Human health.”

(Source: TIMSS 2011 Science Framework, p. 53)

Commonalities between the learning content described in the Curriculum Questionnaire and the TIMSS 2011 Assessment Framework

It is necessary to determine whether the Curriculum Questionnaire and the TIMSS 2011 Assessment Framework deal with the same learning contents. Accordingly, we considered the commonalities between learning content, matching the topic areas described in the Curriculum Questionnaire with the corresponding topic areas described in the TIMSS 2011 Assessment Framework. The Curriculum Questionnaire’s topic areas of “Human Health” and “Changes in Environments” corresponded to the “Human Health” and “Ecosystem” topic areas described by the TIMSS 2011 Assessment Framework. It should be noted that the topic area “Changes in Environments” described in the Curriculum Questionnaire did not correspond to the TIMSS 2011 Assessment Framework’s “Interaction with the Environment” topic, which concerns the association of physical features and behavior patterns in plants and animals with the environment in which they live, and the ways in which plants and animals adapt themselves to particular environments (Mullis et al. 2009).

The topic area of “Human Health” described in the Curriculum Questionnaire includes such learning content as transmission/prevention of communicable diseases, signs of health/illness, diet, and exercise, as shown in Table 1. The topic area of “Human Health” described in the TIMSS 2011 Assessment Framework consists of two topic areas (objectives): “(1) Recognize ways that common communicable diseases (e.g., colds, influenza) are transmitted; identify signs of health or illness and some methods of preventing and treating illness” (Mullis et al. 2009, p. 56); and “(2) Describe ways of staying healthy including eating a balanced diet and regular exercise; identify common food sources (e.g., fruits, vegetables, grains)” (Mullis et al. 2009, p. 56). Keywords shared by the “Human Health” learning content of the Curriculum Questionnaire and the TIMSS 2011 Assessment Framework include “transmission/transmitted,” “prevention/preventing,” “communicable diseases,” “signs of health or illness,” and “diet and exercise.” Based on

these commonalities, the learning content of the “Human Health” topic areas described by the Curriculum Questionnaire and the TIMSS 2011 Assessment Framework are considered to be equivalent.

The topic area of “Changes in Environments” described in the Curriculum Questionnaire includes such learning content as effects of human activity, pollution and its prevention, as shown in Table 1. The learning content of the “Ecosystems” topic area described in the TIMSS 2011 Assessment Framework comprises three objectives, only a third of which is considered directly correlated to “Changes in Environments”: “(3) Explain ways in which human behavior can have positive or negative effects on the environment; provide general descriptions and examples of the effects of pollution on humans, plants, animals, and their environments, and of ways of preventing or reducing pollution” (Mullis et al. 2009, p. 56). Key words and phrases shared by the learning content of the Curriculum Questionnaire’s “Changes in Environments” topic area and described in the TIMSS 2011 Assessment Framework’s “Ecosystems” topic area include “effects of human activity/human behavior can have positive or negative effects,” and “pollution and its prevention/preventing.” Based on these commonalities, the learning content of the “Changes in Environment” topic area described in the Curriculum Questionnaire and the “Ecosystems” topic area described in the TIMSS 2011 Assessment Framework are considered similar to a certain extent.

Students’ performance at the attained curriculum level in relatively less covered topics

As discussed above, we can relate learning content in the “Human Health” and “Changes in Environments” topic areas described by the Curriculum Questionnaire to student performance by way of the percent correct statistics for items under the TIMSS 2011 Assessment Framework. According to TIMSS 2011, the “Human Health” topic area includes 11 items, while the “Ecosystems” topic area includes 14 items.

Table 7 summarizes the percent correct statistics for Japanese fourth grade students for the 11 items of the “Human Health” topic, along with international averages and some item information. With the help of the TCMA data, information was added to Table 7 concerning whether or not the item was covered in the national curriculum. In addition, the items in Table 7 were divided into two groups based on information about which objective of the TIMSS 2011 Assessment Framework each item belongs to.

As shown in Table 7, none of the 11 “Human Health” items were covered in the Japanese primary science curriculum. This is consistent with this study’s findings at the intended curriculum level. Furthermore, Table 7 shows that Japanese fourth graders performed higher than international averages on all 11 items. It can be interpreted that Japanese fourth graders are able to perform higher than the international averages in this topic area. These achieved curriculum results, along with the results from the intended curriculum level, show that in spite of the learning content for “Human Health” receiving less coverage in the national science curriculum, Japanese fourth graders perform better on this learning content than the international averages. It should be noted that the subject of physical education covers some learning content relevant to the topic of “Human Health”.

According to the Course of Study (MEXT 2008), one of the objectives in physical education for grade 4 and grade 5 is “To enable pupils to understand a healthy life and

Table 7 Percent correct statistics for the items of the topic area of "human health" (Source: TIMSS 2011 Item Information Tables, TIMSS 2011 Fourth Grade Almannas and TIMSS 2011 Fourth Grade Test Curriculum Matching Analysis Data)

Item ID	Item type	Cognitive domain	Label	<input type="radio"/> Item not covered in the national science curriculum <input checked="" type="radio"/> Item covered in the national science curriculum	Topic area bullet (objective)	Percent correct: JPN(a)	Percent correct: intl. ave.(b)	(a)-(b)
S041182	CR	Applying	One thing to avoid flu	<input type="radio"/>	Objective 1	92.8	41.5	51.3
S041181	CR	Applying	Influenza passing between people	<input type="radio"/>	Objective 1	78.2	58.5	19.7
S051137	MC	Applying	Why wash hands before eating	<input type="radio"/>	Objective 1	89.7	73.2	16.5
S051024	CR	Applying	Cover your mouth though not sick	<input type="radio"/>	Objective 1	50.8	34.5	16.3
S031325	CR	Applying	Manuel's temperature rise	<input type="radio"/>	Objective 1	47.2	38.9	8.3
S051173	CR	Knowing	Calcium important for growth	<input type="radio"/>	Objective 2	48.6	20.1	28.5
S031340	MC	Knowing	Best source of calcium	<input type="radio"/>	Objective 2	70.4	46.8	23.6
S041301	CR	Applying	Importance of drinking water	<input type="radio"/>	Objective 2	41.2	27.6	13.6
S041183	CR	Applying	Maintaining good physical health	<input type="radio"/>	Objective 2	22.3	17.4	4.9
S041032	CR	Knowing	Staying in sun with no protection	<input type="radio"/>	Objective 2	81.6	76.8	4.8
S041043	CR	Knowing	How to protect teeth from decay	<input type="radio"/>	Objective 2	62.4	59.7	2.7

physical growth and development, and to develop the qualities and abilities for leading healthy and safe everyday lives” (p. 94); additionally, one of its learning content areas is “Health” (p. 96). This could help to explain Japanese fourth graders’ performance on the “Human Health” learning content. Another contributing factor may be the commonly held school activities to raise awareness of health, usually initiated by Japanese primary school nurses. One recommendation for curriculum revision drawn from these results is that learning content relevant to the topic area of “Human Health” need not be preferentially strengthened in the Japanese primary science curriculum.

Table 8 summarizes the percent correct statistics for Japanese fourth grade students for the 14 items of the “Ecosystems” topic, along with the international averages and some item information. Like Tables 7 and 8 includes information about whether or not the item was covered in the national curriculum. The items in Table 8 were divided into three groups based on which of the TIMSS 2011 Assessment Framework’s objectives they belong to. Five items were related to the third objective, whose learning content aligned with the Curriculum Questionnaire’s “Changes in Environments” topic. The five items concerning the third objective are shown in italics for easy reference.

As shown in Table 8, 13 of the 14 “Ecosystems” topic area items were not covered in the Japanese primary science curriculum. None of the five items shown in italics comprising the third objective were covered in the Japanese primary science curriculum. This is again consistent with the study’s findings at the intended curriculum level. Although Japanese fourth graders performed higher than the international averages on nine of the 14 total “Ecosystems” items, this ratio falls when looking only at the five third-objective items, on only two of which Japanese fourth graders performed higher than the international averages. Although this result should be treated with caution due to the limited number of items being discussed here, it can be reasonably interpreted that Japanese fourth graders do not perform higher than the international averages on learning content concerning “Changes in Environments.” These attained curriculum results, along with the results from the intended curriculum level, show that with respect to learning content concerning the “Changes in Environments” topic area, which receives less coverage in the national science curriculum, Japanese fourth graders do not always perform above the international averages. We hypothesize that this could be the result of the lesser coverage given to this learning content in the national curriculum. This hypothesis becomes more persuasive when considering the general trend of Japanese fourth graders’ performance, which is often higher than the international averages in other topics.

A recommendation for curriculum revision drawn from these results is that learning content concerning the topic area of “Changes in Environments” could be strengthened in the Japanese primary science curriculum.

Conclusions

Statistical analysis provided important information at the intended curriculum level, namely that the topic areas of “Changes in Environments” and “Human Health” as described in the Curriculum Questionnaire were relatively less covered by the Japanese primary science curriculum in comparison to the average of the 59 TIMSS 2011 participating countries. By integrating information obtained at the attained curriculum level, this study revealed that despite less coverage of learning content on the topic of “Human

Table 8 Percent correct statistics for the items of the topic area of "ecosystem" (Source: TIMSS 2011 Item Information Tables, TIMSS 2011 Fourth Grade Almannacs and TIMSS 2011 Fourth Grade Test Curriculum Matching Analysis Data)

Item ID	Item type	Cognitive domain	Label	<input type="radio"/> Item not covered in the national science curriculum <input checked="" type="radio"/> Item covered in the national science curriculum	Topic area bullet (objective)	Percent correct: JPN(a)	Percent correct: intl. ave.(b)	(a)-(b)
S051010	CR	Reasoning	Middle tree did not grow as well	<input type="radio"/>	Objective 1	79.6	45.8	33.8
S041178	MC	Knowing	What plants use energy for	<input type="radio"/>	Objective 1	84.9	51.4	33.5
S051194	CR	Reasoning	Test design for plant seedlings	<input type="radio"/>	Objective 1	44.6	21.0	23.6
S041177	CR	Applying	Living things make their own food	<input type="radio"/>	Objective 1	11.2	13.5	-2.3
S051186	CR	Applying	Write in names in food chain	<input type="radio"/>	Objective 2	90.2	63.9	26.3
S041180	MC	Applying	Correct predator-prey relationship	<input type="radio"/>	Objective 2	77.4	57.0	20.4
S051026Z	CR	Reasoning	T/F about predator/prey- DERIVED	<input type="radio"/>	Objective 2	51.6	31.6	20.0
S051007	CR	Applying	Importance of spiders	<input checked="" type="radio"/>	Objective 2	63.1	51.1	12.0
S031356	MC	Knowing	Which animal is a predator	<input type="radio"/>	Objective 2	82.7	87.6	-4.9
S041037	CR	Reasoning	Pos. or neg. effect of activity	<input type="radio"/>	Objective 3	55.5	38.1	17.4
S051168	CR	Applying	Why plastic objects dangerous	<input type="radio"/>	Objective 3	62.3	49.2	13.1
S041039	CR	Reasoning	Better way to travel around town	<input type="radio"/>	Objective 3	45.0	48.2	-3.2
S031390Z	CR	Applying	Animal extinction- DERIVED	<input type="radio"/>	Objective 3	15.6	27.9	-12.3
S051035	MC	Applying	Decrease air pollution in a city	<input type="radio"/>	Objective 3	13.8	34.1	-20.3

Five items related to the third objective, whose learning content aligned with the Curriculum Questionnaire's "Changes in Environments" topic, are shown in italics for easy reference

Health” in the national science curriculum, Japanese fourth graders perform above than the international averages in this area. On the other hand, Japanese fourth graders do not consistently perform above international averages in learning content concerning the topic area of “Changes in Environments.” The study hypothesizes that this may be the result of the lesser coverage of this learning content in the national curriculum.

This study proposes two recommendations for revision of the Japanese primary science curriculum. First, it is not necessary to preferentially strengthen learning content concerning the topic area of “Human Health.” Second, it may be worth strengthening learning content concerning the topic area of “Changes in Environments,” where it is possible to do so.

Although this study adapted TIMSS’s curriculum model, the implemented curriculum level was not dealt with. The next expected step is a discussion of learning content from the perspective of the implemented curriculum, for instance by examining the textbooks of the major publishing companies.

Authors’ contributions

KM contributed to conception, design, and discussion, and drafted the manuscript. YH contributed to design and methods, and performed the statistical analysis. He has been also involved in sequence alignment and drafting of the manuscript. YS made secondary contributions to conception and discussion, participated in coordination, and helped with the manuscript. All authors read and approved the final manuscript.

Authors’ information

Dr. KM is a senior researcher at the National Institute for Educational Policy Research and one of the National Research Coordinators of the TIMSS study for Japan. His research interests include science curriculum development and lesson analysis.

Dr. YH is a senior researcher at the National Institute for Educational Policy Research and the data manager of the TIMSS study for Japan. His research interests include curriculum analysis and educational evaluation and measurement.

Dr. YS is a professor in science education at Kokugakuin University and an emeritus researcher at the National Institute for Educational Policy Research. He is the former National Research Coordinator of the TIMSS study for Japan. His research interests include scientific literacy and active learning in science lessons.

Author details

¹ National Institute for Educational Policy Research, 3-2-2, Kasumigaseki, Chiyoda-ku, Tokyo, Japan. ² Kokugakuin University, 3-22-1, Shin-ishikawa, Aoba-ku, Yokohama, Kanagawa, Japan.

Acknowledgements

This work was partially supported by JSPS KAKENHI Grant Number 25282047 and 25381060. The authors used the following two data sources: TIMSS 2011 Assessment Copyright 2013, International Association for the Evaluation of Educational Achievement (IEA), published by TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College, Chestnut Hill, MA; and International Association for the Evaluation of Educational Achievement (IEA), IEA Secretariat, Amsterdam, The Netherlands.

Competing interests

This work is published by responsibility of the authors. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the institutions with which the authors are affiliated.

Received: 4 November 2015 Accepted: 3 August 2016

Published online: 26 August 2016

References

- Asparouhov, T., & Muthén, B. (2010a). Plausible values for latent variables using Mplus. Technical Report.
- Asparouhov, T., & Muthén, B. (2010b). Bayesian analysis of latent variable models using Mplus (version 4). Technical Report.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81–105.
- Eid, M. (2000). A multitrait-multimethod model with minimal assumptions. *Psychometrika*, *65*(2), 241–261.
- Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*(4), 457–472.
- Kenny, D. A., & Kashy, D. A. (1992). Analysis of the multitrait-multimethod matrix by confirmatory factor analysis. *Psychological Bulletin*, *112*(1), 165–172.

- Marsh, H. W. (1989). Confirmatory factor analysis of multitrait-multimethod data: many problems and a few solutions. *Applied Psychological Measurement*, 13(4), 335–361.
- Martin, M. O., Mullis, I. V. S., Foy, P., & Stanco, G. M. (2012). *TIMSS 2011 international results in science*. Chestnut Hill: TIMSS & PIRLS International Study Center, Boston College.
- Martin, M. O., Mullis, I. V. S., Gonzalez, E. J., & Chrostowski, S. J. (2004). *TIMSS 2003 international science report*. Chestnut Hill: TIMSS & PIRLS International Study Center, Boston College.
- MEXT. (2008). *Chuugakkou gakushuu sidou youryou [Course of Study for Elementary School] (In Japanese)*.
- Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., & Preuschoff, C. (2009). *TIMSS 2011 assessment frameworks*. Chestnut Hill: TIMSS & PIRLS International Study Center, Boston College.
- Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide. Seventh edition*. Los Angeles: Muthén & Muthén.
- Saruta, Y. (2007). Characteristics of the curriculum and scientific literacy in Japan. *NIER Research Bulletin*, 137, 27–45.
- Widaman, K. F. (1985). Hierarchically nested covariance structure models for multitrait-multimethod data. *Applied Psychological Measurement*, 9(1), 1–26.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
