Linking the first- and second-phase IEA studies on mathematics and science

Documentation

Erika Majoros[1,*] Monica Rosén[1]

September 26, 2022

Contents

[1] Department of Education and Special Education, University of Gothenburg, Gothenburg, Sweden

[*] Correspondence: Erika Majoros <erika.majoros@gu.se>

Acknowledgements

We are grateful for the helpful input of Eugenio J. Gonzalez (Educational Testing Service, Princeton, NJ, USA); and we thank Stefan Johansson (Department of Education and Special Education, University of Gothenburg, Gothenburg, Sweden) for sharing his work regarding the first-phase mathematics studies.

Erika Majoros and Monica Rosén are part of the European Training Network OCCAM. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 76500.

The COMPEAT infrastructure project was supported by the Riksbankens Jubileumsfond (grant number In2006-0401:1-IK).

List of Tables

List of Figures

1 Introduction

The present documentation serves to provide technical information about the approach to linking cognitive items of the *first-* and *second-phase* studies on mathematics and science administered by the International Association for the Evaluation of Educational Achievement (IEA). In his review of the history of ILSAs, Gustafsson (2008) identifies two phases of the IEA,), i.e. before and after 1990, when a new organization was set up. We refer to the first-phase studies as follows:

- First International Mathematics Study (FIMS, 1964)
- Six Subject Survey/ First International Science Study (FISS, 1970)
- Second International Mathematics Study (SIMS, 1980)
- Second International Science Study (SISS, 1984)

We refer to the second-phase studies as follows:

- Third International Mathematics and Science Study (TIMSS, 1995)
- Trends in International Mathematics and Science Study (TIMSS, 1999; 2003; 2007; 2011; 2015)

1.1 Data sources

Data and documentation of the old studies were downloaded from the Center for Comparative Analyses of Educational Achievement (COMPEAT) website[1]. Data and documentation of the TIMSS studies were downloaded from the IEA Study Data Repository.

Data management was done with SPSS 25 (IBM Corp., 2017). IRT analyses were performed with the R package mirt (Chalmers, 2012) for the programming language R (R Core Team, 2021).

1.2 Target populations

The definitions of the target populations are shown in Table 1. we can observe that the IEA changed the sampling strategy from age-based to grade-based starting from SIMS.

2 Mathematics cognitive scale

2.1 Bridge items

The number of mathematics test items at each time point and the number of bridge items i.e. those that overlap between the time points are shown in Table 2. We identified 37 items bridging FIMS and SIMS. Out of the 37 items comprising the first bridge, applying the SIMS taxonomy, 15 items are in the arithmetic, 10 are in the algebra, 5 are in the geometry, 2 are in the measurement, and 5 are in the statistics content domain. The 18 bridge items overlapping from SIMS to TIMSS 1995 covered algebra with 3 items, arithmetic with 6 items, geometry with 3 items, measurement with 3 items, and statistics with 3 items.

[1] COMPEAT is an infrastructure project, with the general aim to build databases of international large-scale studies in educational achievement conducted by IEA and the Organisation for Economic Co-operation and Development (OECD) before 2000, hence supporting secondary analyses of these data.

Table 1: Population Definitions of the Respective Studies

	Target Population
FIMS 1964	All pupils who are 13.0-13.11 years old at the date of testing.
FISS 1970	Population 2: All pupils who are 14.0-14.11 years old at the time of testing.
SIMS 1980	Population A: All students in the grade (year level) where the majority have attained the age of 13.00 to 13.11 years by the middle of the school year.
SISS 1984	Population 2: 14-year-old students (corresponds to Population 2 in FISS).
TIMSS 1995	Population 2: all students enrolled in the two adjacent grades that contain the largest proportion of students of age 13 years at the time of testing.
TIMSS 1999	TIMSS in 1999 used the same definition as TIMSS 1995 to identify the target grades, but assessed students in the upper of the two grades only, the eighth grade in most countries.
TIMSS 2003	Population 2: All students enrolled in the upper of the two adjacent grades that contain the largest proportion of 13-year-olds at the time of testing. This grade level was intended to represent eight, counting from the first year of primary or elementary schooling, and was the eighth grade in most countries.
TIMSS 2007	Eighth grade population: this includes all students enrolled in the grade that represents 8 years of formal schooling, counting from the first year of ISCED Level 1, provided that the mean age at the time of testing is at least 13.5 years. For most countries, the target grade should be the eighth grade or its national equivalent.
TIMSS 2011	Same as in 2007
TIMSS 2015	Same as in 2007

Table 2: Item Overlaps Across the Mathematics Assessments

Number of bridge items		37	18	48	79	95	126	128	124
	FIMS	**SIMS**	**T95**	**T99**	**T03**	**T07**	**T11**	**T15**	**T19**
	70	37	9	5	3				
		162	9	6	4				
			141	37	16				
				115	56	21			
					115	74	40		
						120	86	52	
							91	76	46
								96	78
									122
Total number of items	70	199	159	163	194	215	217	224	246

2.2 Recoding process

2.2.1 ID variables

A new set of IDs was recoded/created as follows

- IDSTDY: four digits indicating the year of the assessment (1964; 1980; 1995; 1999; 2003; 2007; 2011; 2015)
- IDCNTRY: country codes in the old studies were recoded to match those in the TIMSS assessments
- IDSCHL, IDCLSS, and IDSDNT: school-, class-, and student IDs were auto-recoded counting from 1 (there were no classes sampled in FIMS so “000” was coded for everyone)
- ID: concatenating IDSTDY, IDCNTRY, IDSCHL, IDCLSS, and IDSDNT

- IDGRD: a variable indicating the number of years of schooling
- IDSEX: variables indicating gender in all studies were recoded with the values “0” for boys and “1” for girls
- BSDAGE: variables indicating age in FIMS and SIMS were recoded from month to years to match those in TIMSS

2.2.2 Weights

Senate weights (with a sum of 500 per country) were applied (stratum weights in SIMS 1980 were rescaled to senate weights). There were no weight variables in the FIMS 1964 datasets, therefore, we have imputed weights in order to weigh each country equally. Individuals within a country were weighed equally.

There was missing data in the weight variables in the following countries participating in SIMS 1980

- Belgium (Flemish): 53 students in three intact schools
- Finland: no weight variables
- France: 17 students from one intact school
- Israel: 26 students from one intact school
- Nigeria: no weight variables

In the case of Finland and Nigeria, we imputed weights similarly to FIMS 1964, in the case of Belgium (Flemish), France, and Israel, we excluded those cases with missing weight values.

2.2.3 Scoring

In FIMS and SIMS, all items were multiple-choice items with a maximum score of 1. Therefore, scoring was conducted as follows

- 0 = incorrect answer
- 1 = correct answer

In the TIMSS studies, multiple-choice items and constructed-response items with a maximum score of 1 were scored as follows

- 0 = incorrect answer
- 1 = correct answer

Constructed-response items with a maximum score of 2 were scored as follows

- 0 = incorrect answer
- 1 = partially correct answer
- 2 = correct answer

Constructed-response items with a maximum score of 2 were scored as follows - 0 = incorrect answer - 1 = minimally correct answer - 2 = partially correct answer - 3 = correct answer

2.2.4 Missing data

There are different types of missing responses in the data and we have recoded them as follows

- Not interpretable/invalid (7, 90, 97) and omitted (9, 99) answers were recoded as incorrect (0)
- Not administered (8, 98) answers were recoded as missing (-99)
- Not reached (6, 96) answers were recoded
  - Missing (-99) for item parameter estimation and
  - Incorrect (0) answers for person scoring

Cases with all missing answers were excluded from the analyses.

The sample sizes in terms of grades in the mathematics assessments are shown in Table 3.

Table 3: Sample Information of the Respective Mathematics Studies

Year	Country	Grade	Sample size	Mean age	SD age
1964	Australia	7	560	13.29	0.29
1964	Australia	8	3081	13.38	0.64
1964	Australia	9	678	13.65	0.23
1964	Belgium	7	327	13.69	2.17
1964	Belgium	8	2590	14.06	1.23
1964	England	8	2456	13.48	0.30
1964	England	9	2778	14.29	0.36
1964	Finland	7	1318	13.78	0.60
1964	France	7	2593	13.53	1.23
1964	France	8	888	14.13	1.60
1964	Germany	7	4473	13.68	0.82
1964	Israel	8	3335	13.97	0.49
1964	Japan	7	2051	13.44	0.28
1964	Netherlands	7	522	13.46	0.92
1964	Netherlands	8	431	14.06	0.70
1964	Scotland	8	2874	13.26	1.43
1964	Scotland	9	4448	13.98	1.89
1964	United States	7	2230	13.44	2.65
1964	United States	8	6551	14.06	2.25
1980	Belgium (Flemish)	8	3066	14.18	0.59
1980	Belgium (French)	8	2062	14.53	0.94
1980	Canada (Britsh Colombia)	8	2554	13.99	0.50
1980	Canada (Ontario)	8	4878	13.39	0.57
1980	England	9	2653	14.13	0.32
1980	Finland	7	4473	13.84	0.40
1980	France	8	8554	14.11	1.04
1980	Hong Kong	7	5520	13.24	0.91
1980	Hungary	8	1754	14.22	0.52
1980	Israel	8	3692	14.03	0.39
1980	Japan	7	7784	13.46	0.29
1980	Luxembourg	8	2061	14.54	0.74
1980	Netherlands	8	5463	14.43	0.66
1980	New Zealand	8	5340	14.02	0.45
1980	Nigeria	9	1463	16.70	3.14
1980	Scotland	9	1334	14.00	0.35
1980	Swaziland (Eswatini)	9	825	15.60	1.87
1980	Sweden	7	3585	13.92	0.35
1980	Thailand	8	3833	14.24	0.75
1980	United States	8	6783	14.13	0.50

2.3 Scale linking method

2.3.1 Stepwise linking of FIMS and SIMS onto the TIMSS reporting scale

First, we performed a concurrent calibration of item parameters pooling all FIMS and SIMS data with fixing the bridge items' parameters to the values reported for TIMSS 1995. These item parameters are reported after a rescaling procedure in the 1999 assessment cycle (Martin & Mullis, 2000).

Then the student abilities were estimated separately for FIMS and SIMS, drawing five PVs per test-taker. To locate the student ability estimates on the TIMSS reporting scale, the original transformation constants used for the reported TIMSS 1995 scaling needed to be applied. Please note that the Swedish FIMS data is missing.

2.3.2 Item types and IRT models

We used the same IRT models as reported in the TIMSS calibration procedures:

- Multiple-choice items: 3PL model
- Constructed response items with a total score of: 2PL model
- Constructed response items with a total score of 2 or 3: generalized partial credit model

2.4 Parameter drift

The delta plot method was applied for the sets of common items, i.e. anchor tests in the mathematics assessments between FIMS, SIMS, and TIMSS 1995. The delta plots are shown in Figures 1, 2, and 3. Two items were flagged for DIF in the first bridge, i.e., from FIMS to SIMS: ys069 and ys133. Two items were flagged for DIF from SIMS to TIMSS 1995: RM1PTI34 and RM1PTI54. We kept these items as unique items for the calibration.

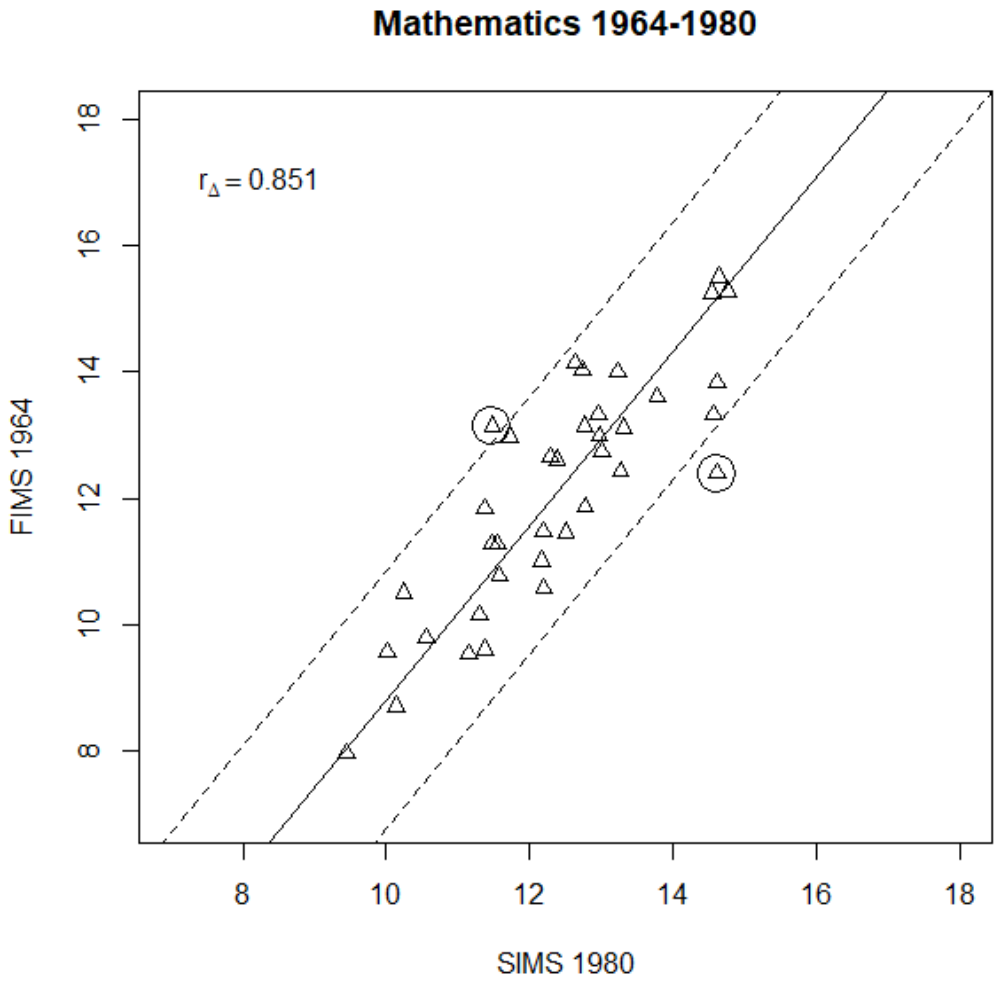

Figure 1: Delta Plots of the Common Items in FIMS and SIMS

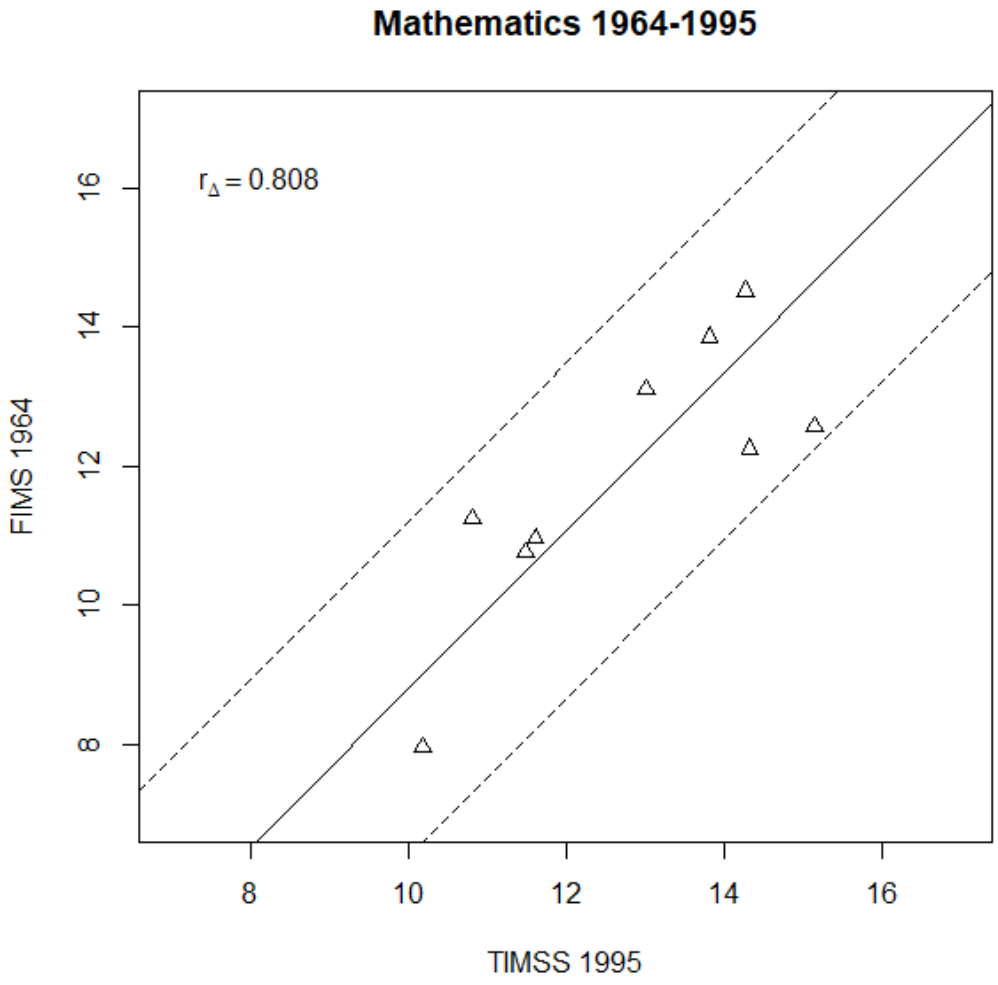

Figure 2: Delta Plots of the Common Items in FIMS and TIMSS 1995

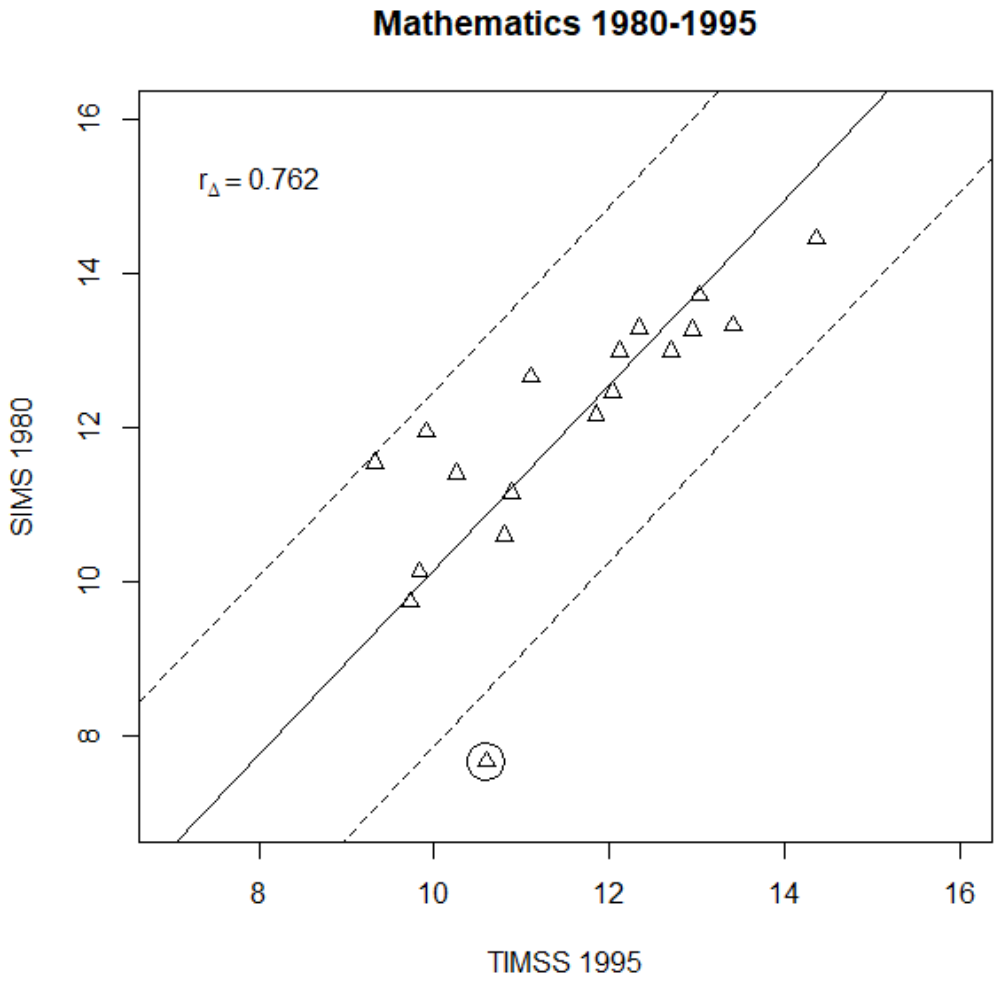

Figure 3: Delta Plots of the Common Items in SIMS and TIMSS 1995

2.5 Item parameter estimates

The item parameter estimates of the FIMS items are shown in Table 4. Common items with TIMSS 1995 were fixed to the reported (Martin & Mullis, 2000) values. In the concurrent calibration, common items with SIMS were assumed and calibrated to have identical estimates.

Table 4: FIMS Item Parameter Estimates

	Slope		Location		Guessing	
Item	Estimate	SE	Estimate	SE	Estimate	SE
RM1PTI1	0.931	0.047	-1.962	0.098	0.000	
RM1PTI2	1.418	0.099	-0.730	0.139	0.146	0.068
RM1PTI3	0.968	0.034	-1.669	0.054	0.000	0.002
RM1PTI4	0.634	0.037	0.266	0.072	0.192	0.024
RM1PTI5	1.054	0.074	0.279	0.094	0.042	0.035
RM1PTI6	0.997	0.051	-0.210	0.093	0.002	0.036
RM1PTI7	1.550	0.057	1.219	0.029	0.000	
RM1PTI8	1.648	0.082	0.610	0.036	0.154	0.014
RM1PTI9	1.292	0.066	0.784	0.040	0.078	0.014
RM1PTI10	1.872	0.162	1.541	0.044	0.154	0.011
RM1PTI11	1.299	0.076	-0.871	0.120	0.004	0.060
RM1PTI12	1.200	0.126	1.114	0.053	0.259	0.016
RM1PTI13	1.967	0.065	0.383	0.020	0.000	
RM1PTI14	0.819	0.115	1.283	0.127	0.162	0.042
RM1PTI15	1.650	0.119	1.117	0.044	0.134	0.016
RM1PTI16	1.589	0.161	1.633	0.053	0.156	0.016
RM1PTI17	1.951	0.145	1.681	0.042	0.135	0.010
RM1PTI18	1.011	0.041	-0.710	0.045	0.000	
RM1PTI19	2.141	0.084	0.977	0.023	0.000	
RM1PTI20	2.172	0.204	1.914	0.051	0.088	0.008
RM1PTI21	2.742	0.619	2.704	0.127	0.266	0.008
RM1PTI22	1.473	0.108	1.413	0.046	0.103	0.013
RM1PTI23	1.360	0.087	-0.250	0.095	0.094	0.043
RM1PTI24	1.471	0.034	-0.489	0.019	0.000	0.001
RM1PTI25	1.535	0.086	-0.702	0.092	0.061	0.049
RM1PTI26	1.258	0.109	0.777	0.045	0.236	0.018
RM1PTI27	1.222	0.060	0.129	0.061	0.078	0.024
RM1PTI28	1.287	0.091	-0.133	0.108	0.122	0.045
RM1PTI29	1.432	0.055	1.389	0.034	0.000	
RM1PTI30	1.209	0.086	1.212	0.049	0.118	0.015
RM1PTI31	0.662	0.028	-0.573	0.086	0.380	0.025
RM1PTI32	1.843	0.115	0.640	0.042	0.133	0.017
RM1PTI33	1.037	0.079	0.097	0.065	0.232	0.027
RM1PTI34	1.509	0.102	0.363	0.061	0.039	0.026
RM1PTI35	0.577	0.034	0.322	0.054	0.000	
RM1PTI36	0.634	0.048	-0.695	0.135	0.198	0.045
RM1PTI37	1.805	0.159	1.206	0.046	0.217	0.016
RM1PTI38	1.185	0.056	0.742	0.026	0.185	0.010
RM1PTI39	1.431	0.098	0.158	0.079	0.158	0.033
RM1PTI40	1.466	0.106	0.604	0.062	0.178	0.024
RM1PTI41	1.294	0.197	2.178	0.092	0.168	0.019

Table 4: FIMS Item Parameter Estimates *(continued)*

	Slope		Location		Guessing	
Item	Estimate	SE	Estimate	SE	Estimate	SE
RM1PTI42	0.998	0.044	0.983	0.037	0.000	
RM1PTI43	3.087	0.591	2.609	0.104	0.157	0.007
RM1PTI44	0.916	0.040	-0.828	0.054	0.000	0.002
RM1PTI45	2.107	0.332	2.356	0.089	0.214	0.010
RM1PTI46	1.593	0.088	0.869	0.038	0.184	0.013
RM1PTI47	1.274	0.090	-0.915	0.179	0.086	0.091
RM1PTI48	1.469	0.089	-0.842	0.116	0.091	0.061
RM1PTI49	0.819	0.024	-0.610	0.031	0.000	0.001
RM1PTI50	1.252	0.080	-0.512	0.126	0.039	0.060
RM1PTI51	1.535	0.054	0.316	0.024	0.000	
RM1PTI52	1.660	0.058	-0.517	0.029	0.000	
RM1PTI53	1.735	0.075	-0.079	0.043	0.136	0.019
RM1PTI54	1.585	0.087	0.312	0.052	0.103	0.023
RM1PTI55	1.077	0.028	0.543	0.023	0.000	0.002
RM1PTI56	1.655	0.095	0.662	0.042	0.092	0.018
RM1PTI57	1.366	0.056	1.271	0.035	0.000	
RM1PTI58	1.292	0.046	-0.184	0.032	0.000	0.007
RM1PTI59	0.900	0.177	2.364	0.148	0.268	0.028
RM1PTI60	0.661	0.148	2.881	0.209	0.135	0.038
RM1PTI61	0.290	0.030	1.336	0.156	0.001	0.009
RM1PTI62	0.935	0.089	0.019	0.200	0.163	0.065
RM1PTI63	2.024	0.197	2.077	0.059	0.140	0.009
RM1PTI64	1.851	0.352	3.053	0.202	0.159	0.008
RM1PTI65	1.416	0.224	2.274	0.117	0.236	0.016
RM1PTI66	0.911	0.154	2.206	0.137	0.177	0.026
RM1PTI67	1.188	0.059	0.566	0.030	0.258	0.012
RM1PTI68	2.683	0.167	0.968	0.027	0.151	0.011
RM1PTI69	1.408	0.068	0.440	0.042	0.108	0.016
RM1PTI70	1.580	0.196	1.711	0.061	0.285	0.017

The item parameter estimates of the FIMS items are shown in Table 5. Common items with TIMSS 1995 were fixed to the reported Martin & Mullis (2000) values. In the concurrent calibration, common items with FIMS were assumed and calibrated to have identical estimates.

Table 5: SIMS Item Parameter Estimates

	Slope		Location		Guessing	
Item	Estimate	SE	Estimate	SE	Estimate	SE
RM1PTI2	1.418	0.099	-0.730	0.139	0.146	0.068
RM1PTI3	0.968	0.034	-1.669	0.054	0.000	0.002
RM1PTI4	0.634	0.037	0.266	0.072	0.192	0.024
RM1PTI5	1.054	0.074	0.279	0.094	0.042	0.035
RM1PTI6	0.997	0.051	-0.210	0.093	0.002	0.036
RM1PTI8	1.648	0.082	0.610	0.036	0.154	0.014
RM1PTI9	1.292	0.066	0.784	0.040	0.078	0.014

Table 5: SIMS Item Parameter Estimates *(continued)*

	Slope		Location		Guessing	
Item	Estimate	SE	Estimate	SE	Estimate	SE
RM1PTI10	1.872	0.162	1.541	0.044	0.154	0.011
RM1PTI11	1.299	0.076	-0.871	0.120	0.004	0.060
RM1PTI12	1.200	0.126	1.114	0.053	0.259	0.016
RM1PTI14	0.819	0.115	1.283	0.127	0.162	0.042
RM1PTI22	1.473	0.108	1.413	0.046	0.103	0.013
RM1PTI23	1.360	0.087	-0.250	0.095	0.094	0.043
RM1PTI24	1.471	0.034	-0.489	0.019	0.000	0.001
RM1PTI25	1.535	0.086	-0.702	0.092	0.061	0.049
RM1PTI26	1.258	0.109	0.777	0.045	0.236	0.018
RM1PTI27	1.222	0.060	0.129	0.061	0.078	0.024
RM1PTI28	1.287	0.091	-0.133	0.108	0.122	0.045
RM1PTI30	1.209	0.086	1.212	0.049	0.118	0.015
RM1PTI31	0.662	0.028	-0.573	0.086	0.380	0.025
RM1PTI32	1.843	0.115	0.640	0.042	0.133	0.017
RM1PTI33	1.037	0.079	0.097	0.065	0.232	0.027
RM1PTI34	1.445	0.068	-0.140	0.054	0.062	0.024
RM1PTI36	0.634	0.048	-0.695	0.135	0.198	0.045
RM1PTI37	1.805	0.159	1.206	0.046	0.217	0.016
RM1PTI38	1.185	0.056	0.742	0.026	0.185	0.010
RM1PTI39	1.431	0.098	0.158	0.079	0.158	0.033
RM1PTI40	1.466	0.106	0.604	0.062	0.178	0.024
RM1PTI46	1.593	0.088	0.869	0.038	0.184	0.013
RM1PTI48	1.469	0.089	-0.842	0.116	0.091	0.061
RM1PTI49	0.819	0.024	-0.610	0.031	0.000	0.001
RM1PTI53	1.735	0.075	-0.079	0.043	0.136	0.019
RM1PTI54	1.382	0.071	1.112	0.024	0.163	0.007
RM1PTI55	1.077	0.028	0.543	0.023	0.000	0.002
RM1PTI66	0.911	0.154	2.206	0.137	0.177	0.026
RM1PTI67	1.188	0.059	0.566	0.030	0.258	0.012
RM1PTI69	1.408	0.068	0.440	0.042	0.108	0.016
ys002	0.781	0.038	0.253	0.048	0.164	0.018
ys004	1.032	0.034	-0.936	0.041	0.000	0.010
ys006	1.053	0.064	0.477	0.067	0.074	0.024
ys007	1.825	0.096	0.046	0.047	0.210	0.020
ys008	0.572	0.023	0.699	0.046	0.000	0.001
ys009	1.849	0.101	0.458	0.037	0.219	0.014
ys012	1.150	0.030	-0.561	0.023	0.000	0.001
ys013	1.478	0.070	0.273	0.042	0.106	0.017
ys014	1.496	0.098	0.254	0.066	0.276	0.023
ys018	1.415	0.074	0.292	0.050	0.124	0.020
ys019	1.384	0.092	0.734	0.053	0.223	0.018
ys020	0.823	0.079	0.351	0.160	0.081	0.050
ys021	0.789	0.086	0.870	0.079	0.246	0.026
ys022	1.190	0.077	-0.442	0.116	0.131	0.047
ys024	1.056	0.080	0.270	0.102	0.135	0.035

Table 5: SIMS Item Parameter Estimates *(continued)*

	Slope		Location		Guessing	
Item	Estimate	SE	Estimate	SE	Estimate	SE
ys025	0.922	0.094	1.379	0.079	0.145	0.024
ys027	1.061	0.105	1.503	0.068	0.171	0.019
ys028	0.723	0.025	-0.307	0.033	0.000	0.001
ys029	1.554	0.109	1.189	0.041	0.125	0.012
ys030	0.760	0.109	-0.669	0.496	0.335	0.115
ys031	0.993	0.036	-0.695	0.063	0.001	0.023
ys032	1.044	0.038	-1.515	0.049	0.000	0.006
ys033	1.330	0.050	-1.784	0.051	0.000	0.012
ys034	1.025	0.041	-1.952	0.064	0.000	0.006
ys036	0.873	0.080	-1.088	0.322	0.203	0.105
ys037	1.801	0.112	1.046	0.034	0.149	0.011
ys039	1.019	0.034	-0.864	0.034	0.000	0.001
ys041	0.871	0.060	-1.927	0.113	0.000	0.006
ys043	1.771	0.109	0.542	0.038	0.095	0.015
ys044	1.227	0.131	-0.033	0.148	0.218	0.053
ys046	1.449	0.121	0.167	0.082	0.178	0.032
ys047	1.562	0.114	-0.021	0.076	0.151	0.032
ys048	0.653	0.128	-0.736	0.747	0.050	0.216
ys049	1.057	0.134	-0.361	0.250	0.173	0.088
ys050	0.902	0.127	-1.320	0.480	0.021	0.199
ys051	1.256	0.280	1.800	0.138	0.184	0.027
ys052	0.746	0.192	1.221	0.284	0.283	0.072
ys053	1.861	0.146	0.071	0.063	0.144	0.027
ys054	1.418	0.144	-0.331	0.143	0.170	0.060
ys055	1.458	0.136	0.901	0.059	0.081	0.020
ys056	1.518	0.163	0.280	0.096	0.162	0.036
ys057	1.173	0.161	-0.095	0.214	0.225	0.074
ys058	1.722	0.203	1.218	0.067	0.201	0.018
ys060	0.920	0.132	-0.403	0.315	0.017	0.118
ys061	1.340	0.169	0.581	0.108	0.138	0.037
ys062	1.774	0.167	0.983	0.054	0.127	0.016
ys063	1.419	0.188	1.048	0.080	0.171	0.027
ys064	1.031	0.078	-1.816	0.110	0.000	0.011
ys065	0.797	0.050	-0.822	0.068	0.000	0.007
ys066	1.655	0.136	-0.275	0.091	0.154	0.041
ys068	1.372	0.062	-1.876	0.057	0.000	0.001
ys069	1.507	0.066	-2.012	0.057	0.000	0.001
ys070	1.371	0.152	0.970	0.075	0.145	0.024
ys071	1.954	0.148	0.521	0.048	0.185	0.019
ys072	1.493	0.153	0.611	0.077	0.183	0.028
ys073	0.954	0.070	-2.186	0.131	0.000	0.009
ys074	1.131	0.126	0.730	0.096	0.075	0.035
ys076	1.112	0.109	1.077	0.066	0.082	0.023
ys077	0.784	0.142	-0.620	0.567	0.185	0.164
ys079	1.041	0.041	-0.087	0.037	0.000	0.008

Table 5: SIMS Item Parameter Estimates *(continued)*

	Slope		Location		Guessing	
Item	Estimate	SE	Estimate	SE	Estimate	SE
ys080	1.214	0.139	-0.267	0.185	0.182	0.070
ys081	1.238	0.126	1.085	0.071	0.047	0.021
ys082	0.819	0.085	0.166	0.168	0.002	0.055
ys083	1.086	0.242	1.069	0.175	0.230	0.055
ys084	1.029	0.147	1.158	0.095	0.097	0.035
ys085	1.236	0.116	-0.484	0.158	0.060	0.070
ys087	1.686	0.216	1.721	0.087	0.150	0.014
ys088	1.201	0.129	0.719	0.088	0.088	0.032
ys089	0.877	0.125	-0.820	0.430	0.097	0.153
ys090	1.229	0.136	0.062	0.142	0.167	0.052
ys091	0.655	0.053	-1.842	0.147	0.001	0.015
ys092	1.658	0.228	0.437	0.107	0.237	0.038
ys093	1.618	0.211	1.364	0.073	0.137	0.018
ys094	1.475	0.244	0.545	0.141	0.297	0.044
ys095	1.037	0.072	-0.806	0.070	0.000	0.010
ys096	1.059	0.141	0.945	0.112	0.113	0.038
ys097	0.668	0.050	0.666	0.071	0.287	0.022
ys098	1.124	0.138	-0.231	0.208	0.127	0.079
ys100	0.956	0.041	0.087	0.040	0.000	0.006
ys101	1.324	0.171	-1.236	0.342	0.314	0.142
ys102	0.645	0.167	0.398	0.587	0.144	0.146
ys103	2.039	0.119	-0.297	0.048	0.111	0.023
ys104	1.141	0.108	-0.396	0.174	0.023	0.076
ys107	1.799	0.145	-0.132	0.076	0.123	0.035
ys109	0.827	0.055	-0.684	0.093	0.194	0.037
ys110	1.809	0.191	0.524	0.072	0.238	0.027
ys112	1.305	0.120	0.352	0.086	0.080	0.034
ys113	1.387	0.150	0.690	0.084	0.173	0.030
ys114	1.458	0.288	1.401	0.119	0.290	0.029
ys117	1.162	0.157	0.571	0.134	0.155	0.047
ys119	1.816	0.187	-0.269	0.103	0.121	0.049
ys121	0.962	0.039	-0.764	0.041	0.000	0.002
ys122	1.231	0.050	-1.011	0.039	0.000	0.001
ys123	1.315	0.161	1.229	0.074	0.105	0.022
ys124	1.831	0.227	1.095	0.071	0.105	0.020
ys125	0.848	0.300	2.545	0.306	0.162	0.043
ys126	1.379	0.133	0.348	0.086	0.066	0.034
ys127	1.164	0.110	0.800	0.071	0.096	0.027
ys129	2.234	0.394	1.935	0.101	0.188	0.012
ys130	1.251	0.164	0.588	0.120	0.205	0.041
ys131	1.343	0.136	0.132	0.109	0.096	0.044
ys132	0.868	0.171	-1.363	0.844	0.243	0.275
ys133	0.913	0.045	-0.564	0.069	0.321	0.027
ys134	0.524	0.048	-0.062	0.123	0.001	0.022
ys135	0.747	0.017	-0.421	0.029	0.062	0.012

Table 5: SIMS Item Parameter Estimates *(continued)*

	Slope		Location		Guessing	
Item	Estimate	SE	Estimate	SE	Estimate	SE
ys136	1.103	0.156	1.474	0.091	0.135	0.029
ys139	1.242	0.126	0.516	0.099	0.110	0.037
ys140	1.731	0.106	0.366	0.042	0.098	0.018
ys141	1.054	0.121	-0.527	0.243	0.038	0.100
ys142	1.178	0.178	1.527	0.099	0.088	0.026
ys143	0.431	0.024	-0.363	0.121	0.301	0.028
ys144	1.238	0.156	0.522	0.123	0.168	0.045
ys145	1.201	0.133	0.255	0.126	0.095	0.048
ys146	1.051	0.062	-0.742	0.059	0.000	0.003
ys147	1.208	0.219	0.993	0.145	0.209	0.047
ys150	1.733	0.206	0.283	0.099	0.162	0.041
ys151	1.478	0.123	1.245	0.049	0.104	0.014
ys152	1.300	0.162	-0.722	0.258	0.173	0.112
ys153	0.823	0.078	-0.354	0.150	0.002	0.045
ys154	0.419	0.060	-0.559	0.210	0.001	0.027
ys156	1.452	0.138	0.464	0.076	0.200	0.028
ys157	1.779	0.196	0.177	0.092	0.105	0.040
ys158	1.283	0.160	-0.325	0.208	0.189	0.082
ys159	1.013	0.180	0.411	0.247	0.301	0.070
ys162	0.791	0.043	-0.120	0.071	0.292	0.025
ys163	1.008	0.161	0.880	0.153	0.147	0.052
ys164	1.712	0.164	-0.589	0.129	0.196	0.063
ys165	1.082	0.042	-0.018	0.033	0.000	0.004
ys166	1.997	0.196	1.413	0.055	0.060	0.011
ys167	0.881	0.039	-1.019	0.051	0.000	0.001
ys168	2.373	0.221	1.862	0.056	0.087	0.006
ys169	1.814	0.340	1.928	0.111	0.228	0.014
ys170	1.026	0.311	2.553	0.293	0.226	0.029
ys171	0.943	0.173	1.966	0.156	0.167	0.028
ys172	1.682	0.851	2.969	0.506	0.199	0.015
ys173	1.968	0.546	2.758	0.267	0.158	0.010
ys174	1.228	0.167	0.726	0.116	0.172	0.043
ys175	1.270	0.080	0.659	0.044	0.011	0.015
ys176	1.065	0.259	2.137	0.181	0.177	0.029
ys177	1.707	0.292	0.201	0.146	0.201	0.061
ys178	0.741	0.274	0.861	0.496	0.156	0.143
ys179	1.864	0.145	0.897	0.040	0.080	0.013
ys180	1.591	0.451	1.661	0.171	0.190	0.031
ys181	1.348	0.064	-0.715	0.040	0.000	0.004
ys182	0.837	0.254	1.004	0.300	0.085	0.104
ys183	1.259	0.246	0.348	0.199	0.121	0.079
ys184	1.681	0.288	0.528	0.115	0.150	0.047
ys185	1.077	0.225	-0.184	0.358	0.124	0.135
ys186	0.832	0.320	-0.264	0.967	0.274	0.255
ys187	1.613	0.239	-0.411	0.185	0.099	0.093

Table 5: SIMS Item Parameter Estimates *(continued)*

	Slope		Location		Guessing	
Item	Estimate	SE	Estimate	SE	Estimate	SE
ys188	0.492	0.080	0.023	0.208	0.001	0.034
ys189	0.930	0.045	-0.282	0.042	0.000	0.001
ys190	1.348	0.068	-0.320	0.044	0.000	0.012
ys191	2.419	0.252	1.138	0.047	0.063	0.014
ys192	1.813	0.293	0.271	0.122	0.195	0.052
ys193	0.433	0.090	-0.012	0.470	0.003	0.091
ys194	1.520	0.437	1.532	0.164	0.232	0.037
ys195	1.695	0.217	0.063	0.104	0.039	0.048
ys196	0.987	0.053	-0.772	0.051	0.000	0.001
ys197	1.002	0.249	0.237	0.351	0.146	0.121
ys198	1.077	0.292	0.641	0.281	0.218	0.093
ys199	0.347	0.079	-1.789	0.489	0.002	0.043

2.6 The scales

The FIMS scale is publicly available on the COMPEAT website following the links below:

- A dataset containing the five plausible scores (FIMSPV1-FIMSPV5) with the new ID variables in SPSS format: PV_FIMS.zip
- A dataset containing the the old ID variables (whole original samples) and the new ID variables in SPSS format: ALL_FIMS_ID_translate.zip
- The above two datasets in R format: FIMS.RData.zip

The SIMS scale is publicly available on the COMPEAT website following this llinks below:

- A dataset containing the five plausible scores (SIMSPV1-SIMSPV5) with the new ID variables in SPSS format: PV_SIMS.zip
- A dataset containing the old ID variables (whole original samples) and the new ID variables in SPSS format: ALL_SIMS_ID_translate.zip
- The above two datasets in R format: SIMS.RData.zip

Please note that cases with extreme grades have been excluded from the linking process. Therefore, they are not scored. Descriptive statistics of the mathematics scales are shown in Table 6.

Table 6: Weighted Statistics of the Mathematics Plausible Scores

Year	Country	Grade	N	Mean PV	Std. error	Std. deviation
1964	Australia	7	560	420.406	3.639	69.747
1964	Australia	8	3081	489.238	1.663	78.490
1964	Australia	9	678	550.831	3.044	76.952
1964	Belgium (Flemish)	7	327	490.706	4.645	72.732
1964	Belgium (Flemish)	8	2590	559.025	1.611	77.301
1964	Finland	7	1318	480.848	2.379	80.049
1964	France	7	2593	515.919	1.450	71.388
1964	France	8	888	592.759	2.684	63.653
1964	Germany	7	4473	550.524	1.144	69.731
1964	Israel	8	3335	576.906	1.449	81.471
1964	Japan	7	2051	553.792	2.231	97.941

Table 6: Weighted Statistics of the Mathematics Plausible Scores *(continued)*

Year	Country	Grade	N	Mean PV	Std. error	Std. deviation
1964	Netherlands	7	522	532.790	3.534	80.264
1964	Netherlands	8	431	551.925	4.207	82.710
1964	United States	7	2230	440.281	1.723	76.752
1964	United States	8	6551	485.323	1.106	83.109
1964	England - GBR	8	2456	531.177	2.142	102.840
1964	England - GBR	9	2778	558.836	2.188	109.192
1964	Scotland - GBR	8	2874	491.084	1.989	95.355
1964	Scotland - GBR	9	4448	503.688	1.557	93.390
1980	Finland	7	4473	467.490	1.676	103.175
1980	France	8	8554	529.992	0.967	74.209
1980	Hong Kong-CHN	7	5520	487.278	1.572	107.267
1980	Hungary	8	1754	550.803	2.479	102.934
1980	Israel	8	3692	513.313	2.053	101.765
1980	Japan	7	7784	565.589	1.314	99.715
1980	Luxemburg	8	2061	506.857	1.814	79.277
1980	Netherlands	8	5463	548.577	1.503	100.880
1980	New Zealand	8	5340	470.915	1.544	103.136
1980	Nigeria	9	1463	432.075	2.271	75.474
1980	Swaziland (Eswatini)	9	825	423.926	2.978	79.393
1980	Sweden	7	3585	436.615	1.760	94.907
1980	Thailand	8	3833	458.755	1.544	91.082
1980	United States	8	6783	476.314	1.380	107.460
1980	England - GBR	9	2653	476.391	2.577	111.859
1980	Scotland - GBR	9	1334	498.979	3.451	117.520
1980	Belgium (French)	8	3066	541.412	1.910	99.251
1980	Bosnia and Herzegovina	8	2062	526.540	2.327	96.396
1980	Canada (Ontario)	8	4878	497.564	1.573	93.120
1980	Canada (Britsh Colombia)	8	2554	467.959	2.815	138.775

3 Science cognitive scale

3.1 Bridge items

The number of items at each time point and the number of bridge items that overlap between the are shown in Table 7.

3.2 Recoding process

3.2.1 ID variables

A new set of IDs was recoded/created as follows

- IDSTDY: four digits indicating the year of the assessment (1970; 1984; 1995)
- IDCNTRY: country codes in the old studies were recoded to match those in the TIMSS assessments
- IDSCHL, IDCLSS, and IDSDNT: school-, class-, and student IDs were auto-recoded counting from 1 (there were no classes sampled in FIMS so “000” was coded for everyone)
- ID: concatenating IDSTDY, IDCNTRY, IDSCHL, IDCLSS, and IDSDNT
- IDGRD: a variable indicating the number of years of schooling

Table 7: Item Overlaps Across the Science Assessments

Number of bridge items		32	13	48	76	102	136	146	153
	FISS	**SISS**	**T95**	**T99**	**T03**	**T07**	**T11**	**T15**	**T19**
	80	32	6	2	1				
		38	7	3	1				
			129	43	22				
				98	52	22			
					123	80	45		
						125	91	52	
							110	94	55
								122	98
									131
Total number of items	80	70	142	146	199	227	246	268	284

- IDSEX: variables indicating gender in all studies were recoded with the values “0” for boys and “1” for girls
- BSDAGE: variables indicating age in FISS and SISS were recoded from month to years to match those in TIMSS

3.2.2 Weights

Senate weights (with a sum of 500 per country) were applied (weights in FISS 1970 and SISS 1984 were rescaled to senate weights).

3.2.3 Scoring

In FISS and SISS, all items were multiple-choice items with a maximum score of 1. Therefore, scoring was conducted as follows

- 0 = incorrect answer
- 1 = correct answer

In TIMSS 1995 studies, multiple-choice items and constructed-response items with a maximum score of 1 were scored as follows

- 0 = incorrect answer
- 1 = correct answer

Constructed-response items with a maximum score of 2 were scored as follows

- 0 = incorrect answer
- 1 = partially correct answer
- 2 = correct answer

3.2.4 Missing data

There are different types of missing responses in the data and we have recoded them as follows

- Not interpretable/invalid (7, 90, 97) and omitted (9, 99) answers were recoded as incorrect (0)
- Not administered (8, 98) answers were recoded as missing (-99)
- Not reached (6, 96) answers were recoded
  - Missing (-99) for item parameter estimation and
  - Incorrect (0) answers for person scoring

Cases with all missing answers were excluded from the analyses.

The sample sizes in terms of grades in the mathematics assessments are shown in Table 8.

Table 8: Sample Information of the Respective Science Studies

Year	Country	Grade	Sample size	Mean age	SD age
1970	Australia	7	11	14.36	0.44
1970	Australia	8	458	14.24	0.24
1970	Australia	9	2680	14.33	0.28
1970	Australia	10	2068	14.57	0.28
1970	Belgium (Flemish)	7	10	15.16	0.33
1970	Belgium (Flemish)	8	48	15.09	0.29
1970	Belgium (Flemish)	9	567	15.02	0.29
1970	Belgium (Flemish)	10	25	15.22	0.24
1970	Belgium (French)	8	199	14.56	0.31
1970	Belgium (French)	9	328	14.56	0.27
1970	Chile	7	176	14.45	0.32
1970	Chile	8	365	14.44	0.30
1970	Chile	9	290	14.60	0.30
1970	Chile	10	55	14.74	0.31
1970	England	7	1	14.67	
1970	England	8	1466	14.30	0.22
1970	England	9	1542	14.77	0.23
1970	England	10	11	14.95	0.36
1970	Finland	7	816	14.35	0.32
1970	Finland	8	1365	14.63	0.24
1970	Finland	9	1	14.92	
1970	Germany	7	207	14.29	0.28
1970	Germany	8	1002	14.36	0.30
1970	Germany	9	947	14.58	0.26
1970	Germany	10	4	14.75	0.15
1970	Hungary	8	4600	14.32	0.24
1970	Hungary	9	2349	14.76	0.10
1970	India	7	782	14.12	0.87
1970	India	8	1058	14.28	0.69
1970	India	9	374	14.53	0.74
1970	India	10	177	14.71	0.66
1970	Iran	7	53	14.21	0.45
1970	Iran	8	1204	14.30	0.43
1970	Iran	9	43	14.31	0.33
1970	Italy	7	18	13.64	0.92
1970	Italy	8	3183	14.49	0.37
1970	Italy	9	3810	14.75	0.42
1970	Italy	10	155	16.47	0.94
1970	Japan	9	1945	14.44	0.29
1970	Netherlands	7	199	14.43	0.32
1970	Netherlands	8	829	14.51	0.29
1970	Netherlands	9	132	14.77	0.23
1970	Netherlands	10	1	15.08	
1970	New Zealand	9	506	14.24	0.24

Table 8: Sample Information of the Respective Science Studies *(continued)*

Year	Country	Grade	Sample size	Mean age	SD age
1970	New Zealand	10	1377	14.53	0.26
1970	Scotland	8	20	14.25	0.19
1970	Scotland	9	883	14.36	0.23
1970	Scotland	10	1056	14.73	0.23
1970	Sweden	7	1201	14.22	0.19
1970	Sweden	8	1207	14.68	0.20
1970	Thailand	7	75	14.26	0.38
1970	Thailand	8	586	14.38	0.37
1970	Thailand	9	977	14.53	0.35
1970	Thailand	10	110	14.73	0.30
1970	United States	7	55	13.87	0.26
1970	United States	8	891	13.85	0.29
1970	United States	9	2391	14.19	0.33
1970	United States	10	7	14.42	0.26
1984	Australia	8	785	14.23	0.23
1984	Australia	9	3500	14.44	0.26
1984	Australia	10	632	14.70	0.21
1984	Canada	8	13	15.02	0.52
1984	Canada	9	7763	15.08	0.57
1984	Canada	10	18	16.06	0.68
1984	China	9	2806	15.64	0.72
1984	England	9	3118	14.24	0.30
1984	Finland	8	2546	14.85	0.34
1984	Ghana	9	2769	16.07	1.68
1984	Hong Kong	8	4973	14.56	0.91
1984	Hungary	8	1334	14.27	0.41
1984	Italy	8	4622	13.93	0.59
1984	Italy	9	1398	14.62	0.24
1984	Japan	9	7610	14.58	0.29
1984	Korea	9	4521	15.02	0.60
1984	Netherlands	9	5025	15.52	1.01
1984	Nigeria	10	804	16.13	0.56
1984	Norway	9	1420	16.81	0.33
1984	Papua New Guinea	10	2193	17.48	1.13
1984	Philippines	9	10871	16.04	1.57
1984	Poland	8	4520	14.96	0.48
1984	Singapore	9	4430	15.23	0.75
1984	Sweden	7	1557	13.86	0.40
1984	Sweden	8	1461	14.82	0.31
1984	Thailand	9	3778	15.36	0.74
1984	United States	9	4477	15.20	0.67
1984	Zimbabwe	9	2395	16.08	1.47

3.3 Scale linking method

3.3.1 Stepwise linking of FISS and SISS onto the TIMSS reporting scale

We followed the same procedure as outlined in 2.3.1:

Linking FISS and SISS to the TIMSS reporting scale was performed in two main steps. First, we performed a concurrent calibration of item parameters pooling all FISS and SISS data with fixing the bridge items' parameters to the values reported for TIMSS 1995.

Then the student abilities were estimated separately for FISS and SISS, drawing five PVs per test-taker. To locate the student ability estimates on the TIMSS reporting scale, the TIMSS 1995 transformation constants were applied.

3.3.2 Item types and IRT models

We used the same IRT models as reported in the TIMSS calibration procedures:

- Multiple-choice items: 3PL model
- Constructed response items with a total score of: 2PL model
- Constructed response items with a total score of 2 or 3: generalized partial credit model

3.4 Parameter drift

We applied the delta plot method on the two bridges between adjacent time points to identify DIF. The delta plots are shown in Figures 4, 5, and 6. Two items were flagged for DIF in the first bridge, i.e., from FISS to SISS. We kept these items as unique items for the calibration.

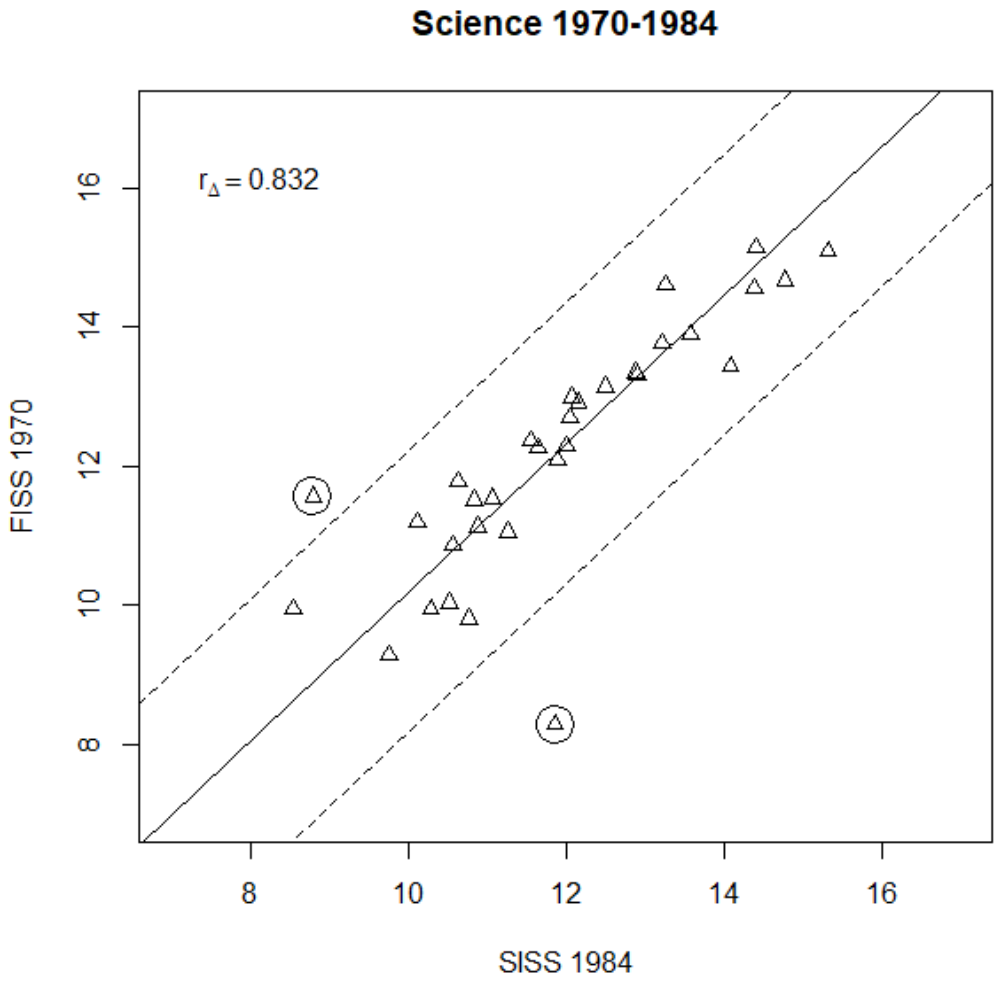

Figure 4: Delta Plot of the Common Items in FISS and SISS

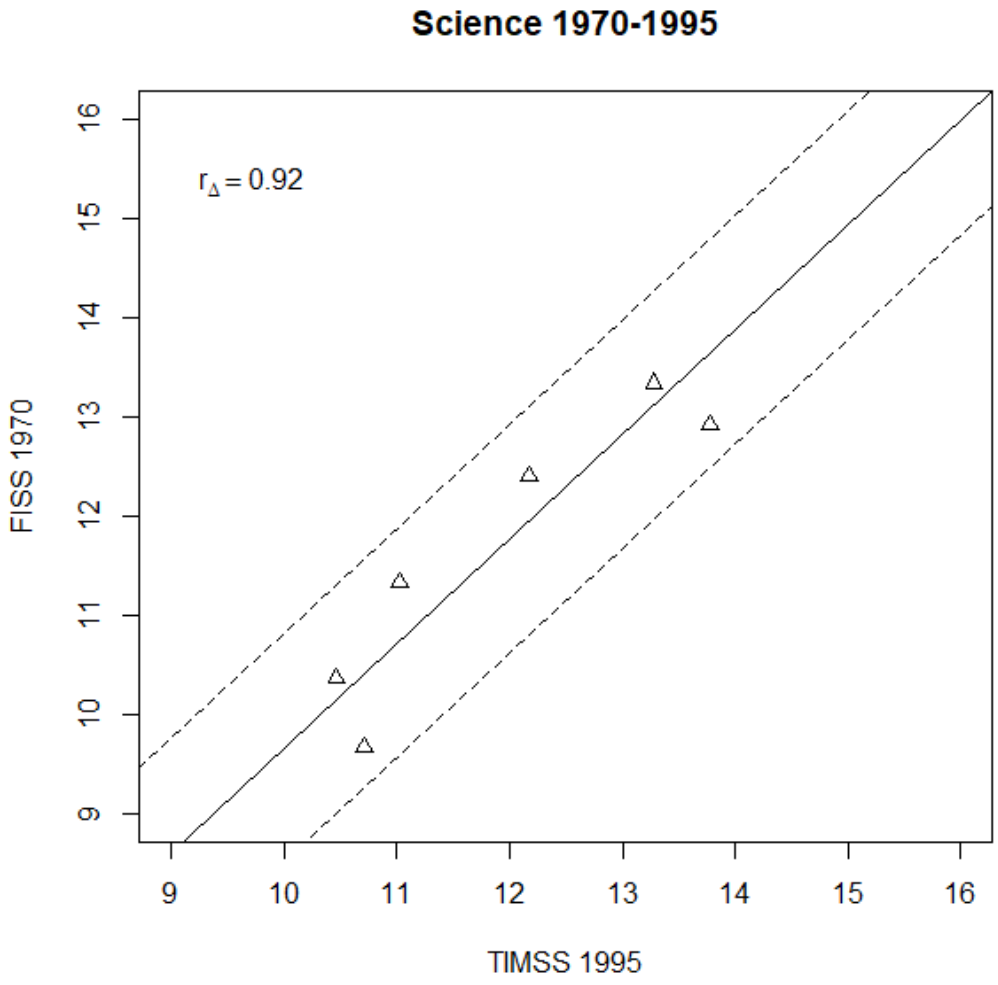

Figure 5: Delta Plot of the Common Items in FISS and TIMSS 1995

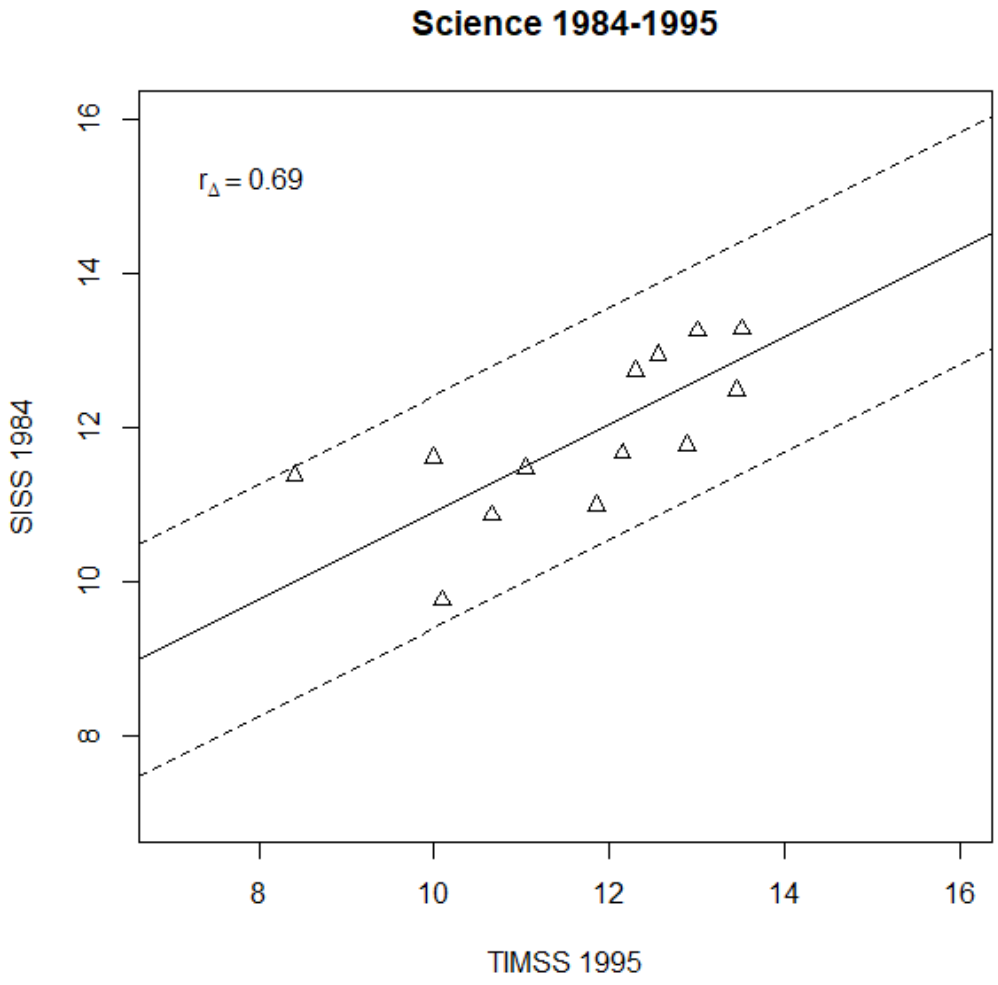

Figure 6: Delta Plot of the Common Items in SISS and TIMSS 1995

3.5 Item parameter estimates

The item parameter estimates of the FISS items are shown in Table 9. Common items with TIMSS 1995 were fixed to the reported (Martin & Mullis, 2000) values. In the concurrent calibration, common items with SISS were assumed and calibrated to have identical estimates.

Table 9: FISS Item Parameter Estimates

	Slope		Location		Guessing	
Item	Estimate	SE	Estimate	SE	Estimate	SE
De2sa01	0.694	0.037	-2.939	0.141	0.001	0.015
De2sa02	0.277	0.015	-0.424	0.061	0.000	0.003
De2sa03	1.385	0.348	2.825	0.238	0.418	0.010
De2sa04	0.565	0.017	-0.770	0.035	0.000	0.001
De2sa05	1.340	0.142	2.391	0.096	0.124	0.008
De2sa06	0.453	0.016	-0.529	0.078	0.209	0.021
De2sa07	0.576	0.059	0.563	0.243	0.010	0.064
De2sa08	0.230	0.024	1.739	0.204	0.000	0.005
De2sa09	0.668	0.028	1.133	0.054	0.000	0.004
De2sa10	1.064	0.194	3.186	0.254	0.144	0.011
De2sa11	0.590	0.026	-0.417	0.050	0.000	0.008
De2sa12	1.362	0.379	3.413	0.379	0.163	0.008
De2sa13	1.126	0.067	1.589	0.038	0.141	0.011
De2sa14	0.486	0.079	1.250	0.159	0.241	0.041
De2sa15	2.207	0.211	2.220	0.067	0.197	0.006
De2sa16	1.154	0.091	1.022	0.062	0.176	0.021
De2sa17	1.678	0.188	2.117	0.072	0.179	0.009
De2sa18	1.563	0.122	1.694	0.050	0.151	0.010
De2sa19	0.640	0.108	2.191	0.129	0.112	0.039
De2sa20	1.318	0.112	1.175	0.057	0.242	0.017
De2sa21	0.784	0.050	-0.719	0.214	0.101	0.069
De2sa22	0.895	0.030	-0.305	0.028	0.000	0.000
De2sa23	1.102	0.075	0.260	0.098	0.318	0.028
De2sa24	2.047	0.123	1.191	0.030	0.165	0.009
De2sa25	0.697	0.081	-0.177	0.333	0.101	0.093
De2sa26	1.507	0.068	1.211	0.028	0.178	0.009
De2sa27	0.795	0.026	-1.963	0.060	0.000	0.005
De2sa28	1.109	0.064	0.152	0.085	0.193	0.029
De2sa29	0.639	0.075	-0.614	0.471	0.211	0.111
De2sa30	1.742	0.141	1.863	0.051	0.149	0.008
De2sa31	0.921	0.041	0.277	0.067	0.035	0.024
De2sa32	1.494	0.298	3.096	0.231	0.156	0.007
De2sa33	1.323	0.847	4.630	1.541	0.207	0.007
De2sa34	2.000	0.133	1.627	0.038	0.128	0.007
De2sa35	1.298	0.153	2.377	0.103	0.168	0.010
De2sa36	0.484	0.025	-0.309	0.050	0.000	0.001
De2sa37	1.873	0.331	3.060	0.196	0.165	0.005
De2sa38	0.281	0.024	0.226	0.085	0.000	0.002
De2sa39	0.783	0.069	0.859	0.111	0.032	0.037
De2sa40	1.364	0.207	2.952	0.187	0.191	0.008
De2sb01	1.525	0.075	0.111	0.056	0.401	0.017

Table 9: FISS Item Parameter Estimates *(continued)*

	Slope		Location		Guessing	
Item	Estimate	SE	Estimate	SE	Estimate	SE
De2sb02	1.233	0.138	1.830	0.076	0.293	0.014
De2sb03	0.800	0.066	-0.311	0.110	0.261	0.040
De2sb04	0.914	0.063	0.757	0.090	0.310	0.023
De2sb05	0.501	0.031	2.966	0.176	0.000	0.001
De2sb06	0.893	0.097	0.728	0.154	0.468	0.030
De2sb07	0.817	0.076	0.353	0.086	0.259	0.030
De2sb08	1.460	0.104	1.609	0.044	0.091	0.010
De2sb09	1.704	0.154	2.159	0.069	0.157	0.007
De2sb10	0.609	0.102	1.811	0.163	0.117	0.050
De2sb11	0.813	0.037	-0.446	0.064	0.216	0.026
De2sb12	1.365	0.104	1.319	0.045	0.144	0.014
De2sb13	0.945	0.042	0.802	0.046	0.052	0.016
De2sb14	0.536	0.066	0.585	0.327	0.037	0.079
De2sb15	-0.088	0.026	-10.950	3.279	0.001	0.017
De2sb16	0.033	0.058	61.943	76.844	0.008	0.180
De2sb17	1.050	0.102	1.610	0.065	0.145	0.019
De2sb18	0.601	0.115	2.130	0.156	0.134	0.048
De2sb19	1.788	0.204	1.856	0.069	0.412	0.009
De2sb20	1.123	0.095	1.331	0.057	0.146	0.018
De2sb21	1.103	0.104	2.257	0.068	0.246	0.010
De2sb22	0.888	0.178	2.109	0.135	0.417	0.025
De2sb23	1.855	0.073	1.352	0.022	0.128	0.005
De2sb24	0.597	0.057	0.557	0.102	0.160	0.032
De2sb25	0.638	0.069	0.886	0.196	0.179	0.047
De2sb26	0.758	0.018	-0.189	0.022	0.000	0.001
De2sb27	1.357	0.092	1.024	0.043	0.109	0.016
De2sb28	1.441	0.114	2.134	0.058	0.188	0.007
De2sb29	1.155	0.056	0.204	0.066	0.179	0.023
De2sb30	1.401	0.066	1.084	0.031	0.152	0.011
De2sb31	1.450	0.102	1.663	0.037	0.275	0.009
De2sb32	1.067	0.167	1.792	0.088	0.340	0.022
De2sb33	1.418	0.083	0.346	0.055	0.135	0.022
De2sb34	1.325	0.102	1.512	0.047	0.119	0.013
De2sb35	2.406	0.234	2.046	0.052	0.153	0.006
De2sb36	1.530	0.196	2.554	0.118	0.188	0.008
De2sb37	1.881	0.152	1.969	0.053	0.132	0.007
De2sb38	0.777	0.110	0.563	0.250	0.193	0.068
De2sb39	1.902	0.135	0.984	0.038	0.244	0.013
De2sb40	0.819	0.111	1.049	0.158	0.220	0.045

The item parameter estimates of the SISS items are shown in Table 10. Common items with TIMSS 1995 were fixed to the reported Martin & Mullis (2000) values. In the concurrent calibration, common items with FISS were assumed and calibrated to have identical estimates.

Table 10: SISS item parameter estimates

	Slope		Location		Guessing	
Item	Estimate	SE	Estimate	SE	Estimate	SE
De2sa01	0.428	0.020	-0.214	0.054	0.000	0.005
De2sa02	0.277	0.015	-0.424	0.061	0.000	0.003
De2sa04	0.565	0.017	-0.770	0.035	0.000	0.001
De2sa06	0.453	0.016	-0.529	0.078	0.209	0.021
De2sa07	0.576	0.059	0.563	0.243	0.010	0.064
De2sa13	1.126	0.067	1.589	0.038	0.141	0.011
De2sa14	0.486	0.079	1.250	0.159	0.241	0.041
De2sa21	0.784	0.050	-0.719	0.214	0.101	0.069
De2sa22	1.101	0.045	-1.301	0.052	0.000	0.002
De2sa23	1.102	0.075	0.260	0.098	0.318	0.028
De2sa26	1.507	0.068	1.211	0.028	0.178	0.009
De2sa27	0.795	0.026	-1.963	0.060	0.000	0.005
De2sa28	1.109	0.064	0.152	0.085	0.193	0.029
De2sa29	0.639	0.075	-0.614	0.471	0.211	0.111
De2sa31	0.921	0.041	0.277	0.067	0.035	0.024
De2sb01	1.525	0.075	0.111	0.056	0.401	0.017
De2sb03	0.800	0.066	-0.311	0.110	0.261	0.040
De2sb04	0.914	0.063	0.757	0.090	0.310	0.023
De2sb06	0.893	0.097	0.728	0.154	0.468	0.030
De2sb07	0.817	0.076	0.353	0.086	0.259	0.030
De2sb11	0.813	0.037	-0.446	0.064	0.216	0.026
De2sb13	0.945	0.042	0.802	0.046	0.052	0.016
De2sb14	0.536	0.066	0.585	0.327	0.037	0.079
De2sb21	1.103	0.104	2.257	0.068	0.246	0.010
De2sb23	1.855	0.073	1.352	0.022	0.128	0.005
De2sb24	0.597	0.057	0.557	0.102	0.160	0.032
De2sb25	0.638	0.069	0.886	0.196	0.179	0.047
De2sb26	0.758	0.018	-0.189	0.022	0.000	0.001
De2sb28	1.441	0.114	2.134	0.058	0.188	0.007
De2sb29	1.155	0.056	0.204	0.066	0.179	0.023
De2sb30	1.401	0.066	1.084	0.031	0.152	0.011
De2sb31	1.450	0.102	1.663	0.037	0.275	0.009
P2A01S	0.701	0.100	-0.504	0.492	0.185	0.128
P2A02S	0.955	0.119	0.409	0.193	0.280	0.054
P2A04S	1.195	0.096	-1.758	0.293	0.072	0.161
P2A05S	1.457	0.063	-0.877	0.058	0.001	0.027
P2A06S	0.795	0.274	3.083	0.360	0.279	0.032
P2A07S	0.708	0.035	-1.339	0.074	0.000	0.003
P2A08S	1.761	0.128	0.788	0.046	0.230	0.017
P2B03S	1.876	0.167	1.537	0.043	0.159	0.011
P2B10S	0.740	0.086	-0.347	0.341	0.117	0.099
P2C01S	0.451	0.032	-0.506	0.172	0.002	0.034
P2C03S	1.381	0.113	0.602	0.073	0.216	0.026
P2C05S	0.561	0.035	0.871	0.060	0.126	0.019
P2C06S	0.346	0.028	-0.182	0.086	0.000	0.004

Table 10: SISS item parameter estimates *(continued)*

	Slope		Location		Guessing	
Item	Estimate	SE	Estimate	SE	Estimate	SE
P2C07S	0.331	0.041	0.820	0.216	0.186	0.042
P2C08S	0.800	0.129	1.571	0.105	0.328	0.022
P2C09S	1.044	0.095	-0.187	0.176	0.180	0.060
P2C10S	1.172	0.091	-0.390	0.146	0.158	0.057
P2D01S	1.138	0.118	0.563	0.123	0.321	0.035
P2D02S	0.919	0.074	0.097	0.079	0.241	0.031
P2D03S	1.215	0.115	0.791	0.088	0.262	0.028
P2D04S	1.593	0.122	0.488	0.065	0.271	0.023
P2D05S	1.617	0.120	0.718	0.053	0.208	0.020
P2D06S	0.540	0.108	-0.257	0.797	0.097	0.181
P2D07S	0.674	0.196	1.897	0.236	0.296	0.066
P2D08S	1.403	0.113	1.703	0.054	0.099	0.011
P2D09S	0.988	0.090	-0.646	0.238	0.154	0.085
P2D10S	0.707	0.039	-0.083	0.098	0.001	0.027
P2M04S	0.914	0.034	-2.418	0.079	0.000	0.002
P2M05S	0.583	0.027	-1.291	0.115	0.179	0.041
P2M06S	0.567	0.060	1.493	0.077	0.344	0.019
P2M07S	1.210	0.077	0.132	0.090	0.295	0.028
P2M11S	1.029	0.088	1.631	0.055	0.153	0.017
P2M15S	1.056	0.060	-1.167	0.172	0.014	0.081
P2M16S	0.717	0.062	-0.617	0.297	0.047	0.092
P2M20S	1.583	0.096	1.283	0.034	0.186	0.011
P2M21S	1.289	0.037	-1.356	0.037	0.000	0.010
P2M27S	0.664	0.063	0.075	0.125	0.239	0.041
P2M28S	0.896	0.032	-2.201	0.071	0.000	0.006

3.6 The scales

The FISS scale is publicly available on the COMPEAT website following this links below:

- A dataset containing the five plausible scores (FISSPV1-FISSPV5) with the new ID variables in SPSS format: PV_FISS.zip
- A dataset containing the the old ID variables (whole original samples) and the new ID variables in SPSS format: ALL_FISS_ID_translate.zip
- The above two datasets in R format: FISS.RData.zip

The SISS scale is publicly available on the COMPEAT website following this links below:

- A dataset containing the five plausible scores (SISSPV1-SISSPV5) with the new ID variables in SPSS format: PV_SISS.zip
- A dataset containing the old ID variables (whole original samples) and the new ID variables in SPSS format: ALL_SISS_ID_translate.zip
- The above two datasets in R format: SISS.RData.zip

Please note that cases with extreme grades have been excluded from the linking process. Therefore, they are not scored. Descriptive statistics of the mathematics scales are shown in Table 11.

Table 11: Weighted Statistics of the Science Plausible Scores

Year	Country	Grade	N	Mean PV	Std. error	Std. deviation
1970	Australia	7	11	475.569	121.809	396.776
1970	Australia	8	458	473.491	4.871	89.008
1970	Australia	9	2680	523.011	1.847	92.034
1970	Australia	10	2068	554.690	2.072	88.608
1970	Chile	7	176	373.326	5.678	68.200
1970	Chile	8	365	410.352	4.140	72.991
1970	Chile	9	290	457.431	5.908	84.638
1970	Chile	10	55	486.974	11.527	80.841
1970	Finland	7	816	490.302	3.179	75.125
1970	Finland	8	1365	505.599	2.262	81.923
1970	Germany	7	207	504.862	6.270	80.563
1970	Germany	8	1002	532.133	2.672	75.310
1970	Germany	9	947	536.699	2.938	80.068
1970	Hungary	8	4600	541.827	1.398	88.319
1970	Hungary	9	2349	545.176	1.980	87.635
1970	India	7	782	370.809	3.366	83.771
1970	India	8	1058	385.233	3.168	88.120
1970	India	9	374	407.703	4.852	89.538
1970	India	10	177	427.913	7.983	88.570
1970	Iran, Islamic Republic of	7	53	364.124	10.598	71.893
1970	Iran, Islamic Republic of	8	1204	406.876	2.651	77.327
1970	Iran, Islamic Republic of	9	43	415.044	14.010	76.298
1970	Italy	7	18	480.886	57.082	240.160
1970	Italy	8	3183	478.166	1.583	83.904
1970	Italy	9	3810	505.105	1.409	82.977
1970	Italy	10	155	597.243	7.884	88.312
1970	Japan	9	1945	560.264	2.311	96.797
1970	Netherlands	7	199	472.486	6.052	73.603
1970	Netherlands	8	829	507.661	2.871	77.311
1970	Netherlands	9	132	540.703	7.725	76.391
1970	New Zealand	9	506	482.169	3.915	81.729
1970	New Zealand	10	1377	539.814	2.334	84.322
1970	Sweden	7	1201	488.064	2.970	80.266
1970	Sweden	8	1207	521.912	2.475	84.570
1970	Thailand	7	75	396.753	9.375	69.484
1970	Thailand	8	586	447.060	3.654	72.104
1970	Thailand	9	977	469.974	2.637	67.376
1970	Thailand	10	110	499.150	8.178	75.509
1970	United States	7	55	421.294	11.716	78.852
1970	United States	8	891	482.279	3.044	88.973
1970	United States	9	2391	516.682	1.768	82.837
1970	United States	10	7	567.206	221.099	583.862
1970	England - GBR	8	1466	491.830	2.491	93.544
1970	England - GBR	9	1542	509.957	2.625	101.273
1970	England - GBR	10	11	541.322	72.751	233.191
1970	Scotland - GBR	8	20	444.790	23.352	98.480

Table 11: Weighted Statistics of the Science Plausible Scores *(continued)*

Year	Country	Grade	N	Mean PV	Std. error	Std. deviation
1970	Scotland - GBR	9	883	497.526	3.501	94.463
1970	Scotland - GBR	10	1056	518.852	3.209	102.885
1970	Belgium (French)	7	10	387.312	23.392	65.726
1970	Belgium (French)	8	48	466.620	11.714	65.670
1970	Belgium (French)	9	567	514.476	3.618	71.760
1970	Belgium (French)	10	25	551.103	12.839	60.362
1970	Bosnia and Herzegovina	8	199	462.243	6.373	74.617
1970	Bosnia and Herzegovina	9	328	484.394	4.812	78.270
1984	Australia	8	785	484.010	4.609	95.064
1984	Australia	9	3500	528.938	2.013	89.588
1984	Australia	10	632	551.939	5.552	90.645
1984	Canada	9	7763	532.872	1.280	86.749
1984	Canada	10	18	493.697	48.837	193.587
1984	China	9	2806	532.540	2.179	91.648
1984	Finland	8	2546	531.653	1.668	77.497
1984	Ghana	9	2769	442.219	2.145	98.915
1984	Hong Kong-CHN	8	4973	497.206	1.224	84.904
1984	Hungary	8	1334	579.586	2.680	93.274
1984	Israel	9	1879	544.328	2.109	89.571
1984	Italy	8	4622	496.618	1.469	90.488
1984	Italy	9	1398	542.883	2.920	87.274
1984	Japan	9	7610	563.779	1.131	95.074
1984	Korea, Republic of	9	4521	524.470	1.372	88.868
1984	Netherlands	9	5025	551.090	1.529	93.361
1984	Nigeria	10	804	422.441	3.527	89.970
1984	Norway	9	1420	524.888	3.251	85.138
1984	Papua New Guinea	10	2193	496.475	1.757	70.605
1984	Philippines	9	10871	403.211	0.922	92.142
1984	Poland	8	4520	525.938	1.570	95.147
1984	Singapore	9	4430	503.350	1.576	94.442
1984	Zimbabwe	9	2395	427.320	2.643	79.384
1984	Sweden	7	1557	512.768	2.514	84.442
1984	Sweden	8	1461	532.250	2.719	92.273
1984	Thailand	9	3778	501.194	1.376	75.760
1984	United States	9	4477	513.728	1.451	95.128
1984	England - GBR	9	3118	509.506	2.644	90.481

References

Chalmers, R. Philip. (2012). mirt: A Multidimensional Item Response Theory Package for the R Environment. *Journal of Statistical Software*, *48*(6), 1–29. https://doi.org/10.18637/jss.v048.i06

Gustafsson, J.-E. (2008). Effects of International Comparative Studies on Educational Quality on the Quality of Educational Research. *European Educational Research Journal*, *7*(1), 1–17. https://doi.org/10.2304/eerj.2008.7.1.1

IBM Corp. (2017). *IBM SPSS Statistics for Windows*. IBM Corp.

Martin, M. O., & Mullis, I. V. S. (2000). TIMSS 1999: an Overview. In M. O. Martin, K. D. Gregory, & S. E. Stemler (Eds.), *TIMSS 1999 Technical Report*. Boston College.

R Core Team. (2021). *R: A language and environment for statistical computing*. https://www.R-project.org/