

**Validity of test score interpretations and cross-cultural comparisons in the First and Second International
Science studies**

Yuriko K. Sosa Paredes¹ and Björn Andersson^{1, 2}

¹Center for Educational Measurement, University of Oslo, Oslo, Norway

²Center for Research on Equality in Education, University of Oslo, Oslo, Norway

Author Note

Yuriko K. Sosa Paredes  <https://orcid.org/0000-0003-1201-4509>

Björn Andersson  <https://orcid.org/0000-0002-9007-2440>

Correspondence concerning this article should be addressed to Björn Andersson, Center for Educational Measurement, Faculty of Educational Sciences, University of Oslo, P.O. Box 1161 Blindern, 0318, Oslo. E-mail: bjorn.andersson@cemo.uio.no

Validity of test score interpretations and cross-cultural comparisons in the First and Second International

Science studies

Supplementary Information (SI)

Table S1

Sample Distribution in FISS and SISS

Educational System	FISS			SISS		
	Sample (original)	Sample (final)	%	Sample (original)	Sample (final)	%
Australia				4259	4259	100.0
Belgium (Flemish)	720	716	99.4			
Belgium (French)	777	767	98.7			
Canada (English)				5104	5104	100.0
Canada (French)				2739	2739	100.0
Chile	1510	1482	98.1			
England	3680	3581	97.3	3748	3748	100.0
FR. Germany	1745	1741	99.8			
Finland	1311	1288	98.2	1600	1600	100.0
Hong Kong				5342	5342	100.0
Hungary	4891	4858	99.3	2596	2590	99.8
India	2731	2679	98.1			
Iran	1708	1641	96.1			
Israel				2351	2351	100.0
Italy	4561	4505	98.8	5156	5156	100.0
Japan	2467	2467	100.0	7924	7924	100.0
South Korea				3489	3489	100.0
Netherlands	1642	1629	99.2			
Nigeria				944	944	100.0
Norway				1305	1305	100.0
Philippines				16851	16851	100.0
Poland				4390	4390	100.0
Singapore				5547	5547	100.0
Scotland	2182	2164	99.2			
Sweden ^a	2041	1984	97.2	1449	1449	100.0
Thailand	1824	1810	99.2			
United States ^b	5550	5483	98.8	2909	2909	100.0

^a Grade 4 in SISS.

^b Phase 1 in SISS.

FISS = First International Science Study, SISS = Second International Science Study.

Table S2*Descriptive Statistics of the Students Scores per Educational System in FISS and SISS*

Educational System	FISS				SISS			
	min	max	mean	sd	min	max	mean	sd
Australia					1	24	13.19	4.45
Belgium (Flemish)	7	37	22.68	5.75				
Belgium (French)	3	34	19.36	5.54				
Canada (English)					1	24	13.75	4.26
Canada (French)					2	24	14.65	4.14
Chile	2	39	15.78	7.16				
England	2	38	20.49	6.78	0	23	11.94	4.39
FR. Germany	3	37	20.88	5.98				
Finland	6	37	22.1	6.59	2	24	15.38	4
Hong Kong					0	24	11.27	4.19
Hungary	2	39	21.57	6.33	2	24	14.68	4.35
India	1	34	15.3	6.83				
Iran	0	27	11.79	4.53				
Israel					1	22	12.47	4.1
Italy	3	39	22.41	7.28	0	24	14.03	4.62
Japan	4	39	25.47	6.08	1	24	15.50	3.94
South Korea					2	24	15.35	4.23
Netherlands	4	38	20.43	6.08				
Nigeria					0	20	7.94	4.01
Norway					1	23	12.75	3.91
Philippines					0	24	9.54	4.47
Poland					0	24	12.46	4.3
Singapore					0	23	10.2	4.2
Scotland	1	38	19.14	6.96				
Sweden ^a	3	38	22.76	5.78	2	24	14.98	3.75
Thailand	1	35	16.51	5.53				
United States ^b	2	39	22.29	7.43	1	24	14.5	4.53

^a Grade 4 in SISS.^b Phase 1 in SISS.

FISS = First International Science Study, SISS = Second International Science Study.

Table S3

*Difficulty Levels and Non-response Ratio in the First
International Science Study Items: E1SA01 to E1SA20*

Item	p	Non-response rate
E1SA01	0.56	0.02
E1SA02	0.67	0.02
E1SA03	0.47	0.04
E1SA04	0.59	0.03
E1SA05	0.53	0.10
E1SA06	0.34	0.05
E1SA07	0.46	0.03
E1SA08	0.41	0.04
E1SA09	0.34	0.02
E1SA10	0.58	0.02
E1SA11	0.69	0.02
E1SA12	0.44	0.06
E1SA13	0.23	0.03
E1SA14	0.51	0.03
E1SA15	0.50	0.04
E1SA16	0.50	0.04
E1SA17	0.63	0.04
E1SA18	0.58	0.04
E1SA19	0.61	0.05
E1SA20	0.36	0.06

Table S4

*Difficulty Levels and Non-response Rate in the First
International Science Study Items: E1SB01 to E1SB20*

Item	p	Non-response rate
E1SB01	0.64	0.02
E1SB02	0.79	0.01
E1SB03	0.73	0.05
E1SB04	0.36	0.02
E1SB05	0.85	0.02
E1SB06	0.48	0.04
E1SB07	0.48	0.04
E1SB08	0.64	0.03
E1SB09	0.34	0.02
E1SB10	0.43	0.04
E1SB11	0.62	0.02
E1SB12	0.33	0.03
E1SB13	0.78	0.02
E1SB14	0.37	0.02
E1SB15	0.49	0.02
E1SB16	0.43	0.04
E1SB17	0.45	0.03
E1SB18	0.38	0.04
E1SB19	0.40	0.04
E1SB20	0.59	0.05

Table S5

Difficulty Levels and Non-response Rate in the Second International Science Study Items

Item	p	Non-response rate
P1M01	0.57	0.01
P1M02	0.49	0.01
P1M03	0.83	0.01
P1M04	0.70	0.01
P1M05	0.58	0.03
P1M06	0.36	0.02
P1M07	0.50	0.01
P1M08	0.49	0.02
P1M09	0.63	0.01
P1M10	0.47	0.01
P1M11	0.49	0.01
P1M12	0.33	0.01
P1M13	0.50	0.02
P1M14	0.43	0.01
P1M15	0.45	0.01
P1M16	0.48	0.02
P1M17	0.51	0.01
P1M18	0.53	0.01
P1M19	0.49	0.02
P1M20	0.52	0.01
P1M21	0.65	0.02
P1M22	0.41	0.02
P1M23	0.43	0.02
P1M24	0.72	0.02

Table S6*Point-biserial Correlations of Items in the First International Science Study*

Educational System	E1SA03	E1SA05	E1SA12	E1SA13	E1SB08	E1SB09
Belgium (Flemish)	0.42	-0.16	0.36	0.20	0.46	0.33
Belgium (French)	0.36	0.43	-0.04	0.09	-0.13	0.20
Chile	0.38	0.47	0.49	0.01	0.46	0.34
England	0.42	0.37	0.43	-0.03	0.48	0.32
FR. Germany	0.38	0.45	-0.07	0.10	0.42	0.26
Finland	0.40	0.45	0.43	0.17	0.42	0.21
Hungary	0.34	0.42	0.43	0.01	0.43	0.47
India	0.31	0.59	0.45	-0.18	0.46	0.43
Iran	-0.04	0.45	0.32	0.04	0.38	-0.04
Italy	0.51	0.49	0.51	0.26	0.47	0.36
Japan	0.42	0.31	0.37	0.15	0.40	0.32
Netherlands	0.35	0.46	0.40	0.13	0.39	0.27
Scotland	0.39	0.35	0.40	-0.03	0.52	0.26
Sweden	0.37	0.45	0.42	0.09	0.46	0.19
Thailand	0.42	0.38	0.31	-0.01	0.42	0.30
United States	0.36	0.42	0.43	-0.05	0.51	0.35

Note. Items with negative correlations are highlighted in bold.

Table S7

Point-biserial Correlations of Items in the Second International Science Study

Educational System	P1M05	P1M19
Australia	0.41	0.42
Canada (English)	0.33	0.36
Canada (French)	0.37	0.39
England	0.45	0.32
Finland	0.40	0.26
Hong Kong	0.45	0.37
Hungary	0.41	0.37
Israel	-0.13	-0.06
Italy	0.47	0.44
Japan	0.28	0.43
South Korea	0.40	0.40
Nigeria	0.52	0.23
Norway	0.44	0.36
Philippines	0.47	0.46
Poland	0.42	0.45
Singapore	0.37	0.26
Sweden ^a	0.38	0.33
United States ^b	0.41	0.45

Note. Items with negative correlations are

highlighted in bold.

^a Grade 4 in SISS.

^b Phase 1 in SISS.

Table S8

Traditional Stepwise Method: Comparing the Baseline Model with Item-Specific Models Across Educational Systems in FISS, EISA01 to EISA20

Model	AIC	BIC	log-likelihood	χ^2	df	p-value
baseline	1323406	1324179	-661611			
EISA01	1322107	1323082	-660938	1346.87	24	0.000
EISA02	1321889	1322864	-660829	1565.18	24	0.000
EISA04	1322589	1323564	-661178	865.40	24	0.000
EISA06	1321891	1322866	-660830	1563.21	24	0.000
EISA07	1321966	1322941	-660867	1487.94	24	0.000
EISA08	1322736	1323711	-661252	718.17	24	0.000
EISA09	1322138	1323113	-660953	1316.12	24	0.000
EISA10	1321781	1322756	-660774	1673.31	24	0.000
EISA11	1319383	1320358	-659576	4071.07	24	0.000
EISA14	1323050	1324025	-661409	404.51	24	0.000
EISA15	1321972	1322947	-660870	1481.79	24	0.000
EISA16	1322689	1323664	-661228	765.43	24	0.000
EISA17	1322476	1323451	-661122	978.32	24	0.000
EISA18	1322088	1323063	-660928	1366.31	24	0.000
EISA19	1322374	1323349	-661071	1079.93	24	0.000
EISA20	1322512	1323487	-661140	942.16	24	0.000

FISS = First International Science Study.

Table S9

Traditional Stepwise Method: Comparing the Baseline Model with Item-Specific Models Across Educational Systems in FISS, E1SB01 to E1SB20

Model	AIC	BIC	log-likelihood	χ^2	df	p-value
baseline	1323406	1324179	-661611			
E1SB01	1322287	1323262	-661028	1167.04	24	0.000
E1SB02	1322478	1323453	-661123	976.60	24	0.000
E1SB03	1322534	1323509	-661151	919.81	24	0.000
E1SB04	1322469	1323444	-661119	985.06	24	0.000
E1SB05	1322280	1323254	-661024	1174.78	24	0.000
E1SB06	1322157	1323131	-660962	1297.77	24	0.000
E1SB07	1322482	1323457	-661125	972.49	24	0.000
E1SB10	1322368	1323343	-661068	1086.15	24	0.000
E1SB11	1320647	1321622	-660208	2806.87	24	0.000
E1SB12	1322757	1323732	-661263	696.96	24	0.000
E1SB13	1321612	1322587	-660690	1842.33	24	0.000
E1SB14	1320794	1321769	-660281	2660.01	24	0.000
E1SB15	1323081	1324056	-661425	373.02	24	0.000
E1SB16	1321611	1322586	-660690	1843.14	24	0.000
E1SB17	1321414	1322389	-660591	2040.14	24	0.000
E1SB18	1321695	1322669	-660731	1759.72	24	0.000
E1SB19	1322639	1323614	-661204	814.99	24	0.000
E1SB20	1321894	1322868	-660831	1560.68	24	0.000

FISS = First International Science Study.

Table S10

Traditional Stepwise Method: Comparing the Baseline Model with Item-Specific Models Across Educational Systems in SISS

Model	AIC	BIC	log-likelihood	χ^2	df	p-value
baseline	2046065	2046768	-1022956			
P1M01	2042305	2043304	-1021045	3823.62	32	0.000
P1M02	2044229	2045228	-1022007	1899.82	32	0.000
P1M03	2044517	2045516	-1022150	1612.15	32	0.000
P1M04	2045172	2046171	-1022478	956.68	32	0.000
P1M06	2043166	2044165	-1021475	2962.56	32	0.000
P1M07	2038953	2039952	-1019369	7175.92	32	0.000
P1M08	2043396	2044395	-1021590	2733.22	32	0.000
P1M09	2044515	2045513	-1022149	1614.42	32	0.000
P1M10	2043593	2044592	-1021688	2536.09	32	0.000
P1M11	2043066	2044065	-1021425	3063.00	32	0.000
P1M12	2042387	2043385	-1021085	3742.36	32	0.000
P1M13	2042758	2043756	-1021271	3371.46	32	0.000
P1M14	2044157	2045155	-1021970	1972.32	32	0.000
P1M15	2043751	2044750	-1021767	2378.01	32	0.000
P1M16	2043684	2044683	-1021734	2444.80	32	0.000
P1M17	2043888	2044887	-1021836	2240.75	32	0.000
P1M18	2044002	2045001	-1021893	2126.69	32	0.000
P1M20	2043508	2044507	-1021646	2620.52	32	0.000
P1M21	2043379	2044378	-1021582	2749.95	32	0.000
P1M22	2041596	2042595	-1020690	4532.60	32	0.000
P1M23	2043038	2044036	-1021411	3091.39	32	0.000
P1M24	2042010	2043009	-1020897	4118.92	32	0.000

SISS = Second International Science Study.

Table S11

The Alignment Method: Approximate Measurement (Non-) Invariance for Intercepts and Slopes in FISS Items, E1SA01 to E1SA20

Intercepts													
E1SA01	BEFL	BEFR	(EN)	FRDE	FI	HU	(IT)	JP	NL	SCL	(SE)	(TH)	(US)
E1SA02	BEFL	(BEFR)	EN	FRDE	(FI)	HU	(IT)	(JP)	(NL)	SCL	SE	(TH)	US
E1SA04	BEFL	(BEFR)	(EN)	(FRDE)	FI	HU	IT	(JP)	NL	SCL	SE	(TH)	(US)
E1SA06	BEFL	(BEFR)	EN	(FRDE)	FI	(HU)	IT	(JP)	NL	SCL	SE	(TH)	(US)
E1SA07	BEFL	(BEFR)	EN	FRDE	FI	(HU)	(IT)	JP	(NL)	SCL	SE	(TH)	US
E1SA08	BEFL	BEFR	(EN)	(FRDE)	(FI)	HU	IT	JP	NL	(SCL)	SE	TH	(US)
E1SA09	BEFL	BEFR	EN	FRDE	FI	HU	(IT)	JP	(NL)	SCL	SE	(TH)	US
E1SA10	(BEFL)	(BEFR)	EN	(FRDE)	(FI)	(HU)	IT	JP	NL	SCL	(SE)	(TH)	(US)
E1SA11	BEFL	(BEFR)	(EN)	(FRDE)	FI	HU	(IT)	(JP)	NL	(SCL)	SE	(TH)	(US)
E1SA14	BEFL	BEFR	EN	(FRDE)	(FI)	(HU)	IT	JP	NL	SCL	(SE)	TH	US
E1SA15	BEFL	BEFR	EN	FRDE	FI	(HU)	IT	(JP)	NL	SCL	(SE)	(TH)	(US)
E1SA16	BEFL	BEFR	EN	(FRDE)	FI	HU	IT	(JP)	(NL)	SCL	(SE)	TH	(US)
E1SA17	(BEFL)	BEFR	EN	(FRDE)	(FI)	(HU)	(IT)	JP	NL	SCL	SE	TH	US
E1SA18	(BEFL)	BEFR	(EN)	(FRDE)	(FI)	(HU)	IT	JP	(NL)	(SCL)	SE	(TH)	(US)
E1SA19	BEFL	BEFR	EN	FRDE	(FI)	(HU)	IT	(JP)	NL	(SCL)	SE	TH	US
E1SA20	BEFL	BEFR	EN	FRDE	FI	HU	(IT)	JP	NL	SCL	SE	TH	US
Slopes													
E1SA01	BEFL	BEFR	(EN)	FRDE	FI	HU	IT	JP	NL	(SCL)	SE	(TH)	US
E1SA02	BEFL	(BEFR)	(EN)	FRDE	FI	HU	IT	JP	NL	SCL	SE	(TH)	US
E1SA04	BEFL	BEFR	(EN)	FRDE	FI	(HU)	IT	(JP)	NL	SCL	SE	(TH)	(US)
E1SA06	(BEFL)	BEFR	EN	FRDE	FI	HU	(IT)	JP	NL	SCL	SE	TH	(US)
E1SA07	BEFL	BEFR	EN	FRDE	FI	HU	(IT)	JP	NL	SCL	(SE)	(TH)	US
E1SA08	(BEFL)	BEFR	EN	FRDE	FI	(HU)	(IT)	JP	NL	SCL	SE	TH	US
E1SA09	BEFL	BEFR	(EN)	FRDE	FI	HU	(IT)	JP	NL	(SCL)	SE	TH	US
E1SA10	BEFL	BEFR	EN	FRDE	FI	HU	(IT)	JP	(NL)	SCL	SE	TH	(US)
E1SA11	BEFL	BEFR	EN	(FRDE)	FI	HU	IT	JP	NL	SCL	SE	(TH)	US
E1SA14	BEFL	BEFR	EN	FRDE	FI	HU	IT	JP	NL	SCL	SE	TH	(US)
E1SA15	BEFL	BEFR	EN	FRDE	FI	HU	(IT)	JP	NL	SCL	SE	TH	US
E1SA16	BEFL	BEFR	EN	FRDE	FI	(HU)	IT	(JP)	NL	SCL	SE	TH	US
E1SA17	BEFL	BEFR	EN	FRDE	FI	HU	(IT)	JP	NL	SCL	SE	TH	US
E1SA18	BEFL	BEFR	EN	FRDE	(FI)	HU	IT	JP	(NL)	SCL	(SE)	TH	(US)
E1SA19	(BEFL)	BEFR	EN	FRDE	FI	HU	IT	JP	(NL)	SCL	SE	TH	US
E1SA20	BEFL	BEFR	EN	FRDE	FI	(HU)	(IT)	(JP)	NL	SCL	SE	TH	US

FISS = First International Science Study, BEFL = Belgium Flemish, BEFR = Belgium French, EN = England, FRDE = FR.

Germany, FI = Finland, HU = Hungary, IT = Italy, JP = Japan, NL = Netherlands, SCL = Scotland, SE = Sweden, TH = Thailand,

US = United States

Table S12

The Alignment Method: Approximate Measurement (Non-) Invariance for Intercepts and Slopes in FISS Items, E1SB01 to E1SB20

Intercepts													
E1SB01	BEFL	BEFR	(EN)	FRDE	FI	(HU)	(IT)	JP	(NL)	SCL	SE	(TH)	US
E1SB02	BEFL	BEFR	EN	FRDE	(FI)	(HU)	(IT)	(JP)	(NL)	SCL	SE	(TH)	(US)
E1SB03	BEFL	BEFR	EN	FRDE	FI	(HU)	(IT)	JP	NL	SCL	SE	(TH)	(US)
E1SB04	BEFL	(BEFR)	EN	(FRDE)	(FI)	HU	(IT)	JP	(NL)	SCL	SE	TH	(US)
E1SB05	BEFL	(BEFR)	EN	(FRDE)	FI	(HU)	(IT)	(JP)	NL	SCL	SE	(TH)	(US)
E1SB06	BEFL	BEFR	(EN)	FRDE	FI	HU	IT	(JP)	NL	(SCL)	SE	(TH)	(US)
E1SB07	BEFL	BEFR	(EN)	FRDE	FI	(HU)	IT	JP	(NL)	(SCL)	SE	(TH)	(US)
E1SB10	BEFL	BEFR	(EN)	FRDE	FI	(HU)	IT	JP	NL	(SCL)	SE	(TH)	US
E1SB11	(BEFL)	(BEFR)	EN	FRDE	FI	(HU)	IT	(JP)	NL	SCL	(SE)	(TH)	(US)
E1SB12	BEFL	BEFR	EN	FRDE	FI	(HU)	IT	(JP)	NL	SCL	(SE)	(TH)	US
E1SB13	BEFL	(BEFR)	(EN)	FRDE	FI	HU	(IT)	(JP)	(NL)	(SCL)	SE	(TH)	(US)
E1SB14	BEFL	BEFR	EN	FRDE	FI	(HU)	IT	(JP)	(NL)	(SCL)	(SE)	(TH)	(US)
E1SB15	BEFL	(BEFR)	EN	FRDE	FI	(HU)	(IT)	JP	NL	SCL	(SE)	TH	US
E1SB16	BEFL	BEFR	EN	FRDE	(FI)	(HU)	(IT)	JP	(NL)	SCL	(SE)	(TH)	US
E1SB17	BEFL	(BEFR)	EN	FRDE	FI	HU	IT	(JP)	(NL)	(SCL)	(SE)	(TH)	(US)
E1SB18	BEFL	BEFR	EN	(FRDE)	FI	HU	(IT)	(JP)	NL	SCL	(SE)	(TH)	US
E1SB19	BEFL	BEFR	EN	(FRDE)	FI	HU	IT	JP	(NL)	SCL	SE	(TH)	(US)
E1SB20	BEFL	BEFR	(EN)	FRDE	FI	(HU)	(IT)	(JP)	NL	(SCL)	(SE)	TH	US
Slopes													
E1SB01	BEFL	BEFR	EN	FRDE	FI	HU	IT	JP	(NL)	SCL	(SE)	TH	(US)
E1SB02	BEFL	BEFR	EN	FRDE	FI	HU	IT	JP	NL	SCL	SE	TH	US
E1SB03	BEFL	BEFR	EN	FRDE	(FI)	HU	IT	JP	NL	SCL	SE	TH	US
E1SB04	BEFL	BEFR	EN	FRDE	FI	HU	(IT)	JP	NL	SCL	(SE)	TH	US
E1SB05	BEFL	BEFR	EN	FRDE	FI	HU	(IT)	JP	NL	SCL	SE	TH	US
E1SB06	BEFL	BEFR	(EN)	FRDE	FI	HU	IT	JP	(NL)	(SCL)	SE	(TH)	US
E1SB07	BEFL	BEFR	EN	(FRDE)	FI	HU	(IT)	JP	NL	SCL	SE	TH	US
E1SB10	BEFL	BEFR	EN	(FRDE)	FI	HU	IT	JP	NL	SCL	SE	TH	US
E1SB11	BEFL	BEFR	EN	FRDE	FI	HU	(IT)	(JP)	NL	SCL	SE	(TH)	(US)
E1SB12	BEFL	BEFR	EN	FRDE	FI	HU	IT	JP	(NL)	SCL	SE	(TH)	US
E1SB13	BEFL	BEFR	EN	FRDE	FI	(HU)	IT	JP	NL	SCL	SE	(TH)	US
E1SB14	BEFL	BEFR	EN	FRDE	FI	HU	IT	(JP)	NL	SCL	SE	(TH)	US
E1SB15	BEFL	BEFR	EN	FRDE	FI	(HU)	IT	JP	NL	SCL	SE	(TH)	(US)
E1SB16	BEFL	BEFR	EN	FRDE	FI	(HU)	IT	JP	NL	SCL	SE	TH	US
E1SB17	BEFL	(BEFR)	(EN)	FRDE	FI	HU	IT	JP	(NL)	SCL	(SE)	TH	US
E1SB18	BEFL	BEFR	EN	FRDE	FI	(HU)	(IT)	(JP)	NL	SCL	SE	(TH)	US
E1SB19	BEFL	BEFR	(EN)	FRDE	FI	HU	(IT)	JP	NL	SCL	SE	(TH)	US
E1SB20	BEFL	BEFR	EN	FRDE	FI	HU	IT	JP	NL	SCL	SE	TH	(US)

FISS = First International Science Study, BEFL = Belgium Flemish, BEFR = Belgium French, EN = England, FRDE = FR.

Germany, FI = Finland, HU = Hungary, IT = Italy, JP = Japan, NL = Netherlands, SCL = Scotland, SE = Sweden, TH = Thailand,

US = United States

Table S13*The Alignment Method: Approximate Measurement (Non-) Invariance for Intercepts and Slopes in SISS Items*

Intercepts																	
P1M01	(AU)	(CAFR)	CAEN	EN	FI	(HK)	HU	(IL)	(IT)	JP	KR	(NO)	(PH)	(PL)	SG	SE	US
P1M02	(AU)	CAFR	(CAEN)	(EN)	FI	(HK)	HU	IL	(IT)	(JP)	KR	NO	(PH)	PL	SG	SE	US
P1M03	(AU)	CAFR	CAEN	(EN)	FI	(HK)	HU	IL	(IT)	JP	(KR)	NO	(PH)	PL	(SG)	SE	US
P1M04	AU	CAFR	CAEN	EN	FI	HK	HU	IL	IT	(JP)	KR	NO	PH	(PL)	SG	SE	(US)
P1M06	(AU)	CAFR	CAEN	(EN)	FI	HK	(HU)	(IL)	IT	(JP)	KR	NO	PH	PL	SG	SE	US
P1M07	(AU)	CAFR	CAEN	EN	FI	HK	(HU)	(IL)	IT	JP	(KR)	NO	(PH)	(PL)	SG	SE	US
P1M08	AU	CAFR	(CAEN)	(EN)	(FI)	(HK)	HU	(IL)	(IT)	JP	KR	NO	PH	(PL)	(SG)	SE	US
P1M09	AU	(CAFR)	CAEN	EN	FI	HK	HU	(IL)	IT	JP	(KR)	NO	PH	(PL)	(SG)	SE	US
P1M10	AU	(CAFR)	CAEN	EN	FI	(HK)	(HU)	(IL)	IT	(JP)	KR	(NO)	(PH)	PL	(SG)	(SE)	US
P1M11	AU	CAFR	(CAEN)	EN	FI	HK	HU	(IL)	IT	JP	KR	(NO)	(PH)	PL	(SG)	SE	US
P1M12	AU	CAFR	CAEN	(EN)	(FI)	(HK)	HU	(IL)	IT	(JP)	(KR)	(NO)	(PH)	(PL)	SG	SE	US
P1M13	AU	(CAFR)	CAEN	EN	FI	HK	(HU)	(IL)	IT	(JP)	(KR)	NO	(PH)	(PL)	SG	SE	US
P1M14	AU	CAFR	(CAEN)	EN	FI	HK	HU	IL	(IT)	JP	KR	NO	(PH)	PL	SG	SE	(US)
P1M15	(AU)	CAFR	CAEN	EN	FI	HK	HU	IL	IT	JP	KR	NO	PH	(PL)	(SG)	SE	US
P1M16	AU	CAFR	CAEN	EN	FI	HK	HU	(IL)	(IT)	(JP)	(KR)	(NO)	(PH)	(PL)	SG	SE	US
P1M17	AU	CAFR	CAEN	EN	(FI)	(HK)	HU	IL	IT	(JP)	KR	NO	PH	PL	SG	SE	(US)
P1M18	AU	CAFR	CAEN	EN	FI	(HK)	(HU)	IL	IT	JP	(KR)	(NO)	(PH)	(PL)	(SG)	SE	US
P1M20	AU	(CAFR)	CAEN	(EN)	FI	(HK)	HU	(IL)	(IT)	(JP)	(KR)	(NO)	(PH)	PL	SG	SE	US
P1M21	AU	CAFR	CAEN	(EN)	FI	HK	HU	IL	IT	JP	KR	NO	PH	(PL)	(SG)	SE	US
P1M22	(AU)	(CAFR)	(CAEN)	(EN)	FI	HK	HU	IL	(IT)	(JP)	(KR)	(NO)	(PH)	(PL)	(SG)	SE	(US)
P1M23	(AU)	CAFR	CAEN	(EN)	FI	HK	HU	IL	(IT)	JP	KR	(NO)	PH	(PL)	(SG)	SE	US
P1M24	AU	CAFR	CAEN	(EN)	FI	(HK)	(HU)	IL	IT	(JP)	KR	NO	(PH)	PL	(SG)	SE	US
Slopes																	
P1M01	AU	CAFR	CAEN	(EN)	FI	(HK)	HU	IL	IT	JP	(KR)	NO	(PH)	PL	(SG)	SE	US
P1M02	AU	CAFR	CAEN	EN	FI	HK	HU	IL	IT	JP	KR	NO	PH	(PL)	SG	SE	US
P1M03	AU	CAFR	CAEN	EN	FI	(HK)	HU	(IL)	IT	JP	KR	NO	(PH)	PL	(SG)	SE	US
P1M04	AU	CAFR	CAEN	EN	FI	HK	HU	IL	IT	JP	KR	NO	(PH)	PL	SG	SE	(US)
P1M06	(AU)	CAFR	CAEN	EN	FI	HK	HU	IL	IT	JP	KR	NO	PH	PL	(SG)	SE	(US)
P1M07	AU	CAFR	CAEN	EN	FI	HK	(HU)	IL	(IT)	(JP)	(KR)	NO	(PH)	PL	(SG)	SE	US
P1M08	AU	CAFR	CAEN	EN	FI	(HK)	(HU)	IL	IT	JP	KR	NO	(PH)	PL	(SG)	SE	US
P1M09	AU	CAFR	CAEN	EN	FI	HK	HU	(IL)	IT	JP	(KR)	NO	(PH)	PL	(SG)	SE	US
P1M10	AU	CAFR	CAEN	EN	FI	HK	(HU)	(IL)	(IT)	(JP)	KR	NO	PH	PL	SG	SE	US
P1M11	AU	CAFR	CAEN	EN	FI	HK	(HU)	(IL)	(IT)	JP	KR	NO	(PH)	PL	(SG)	(SE)	US
P1M12	AU	CAFR	CAEN	EN	FI	HK	(HU)	(IL)	(IT)	(JP)	KR	NO	(PH)	(PL)	SG	SE	US
P1M13	AU	CAFR	CAEN	EN	FI	(HK)	HU	IL	(IT)	(JP)	KR	(NO)	PH	(PL)	SG	SE	US
P1M14	AU	CAFR	CAEN	(EN)	FI	(HK)	HU	IL	IT	JP	KR	NO	(PH)	PL	SG	(SE)	US
P1M15	AU	CAFR	CAEN	EN	FI	HK	HU	IL	(IT)	(JP)	KR	NO	(PH)	PL	(SG)	SE	US
P1M16	AU	CAFR	CAEN	EN	FI	HK	HU	IL	IT	JP	(KR)	NO	(PH)	PL	(SG)	SE	US
P1M17	AU	CAFR	CAEN	EN	(FI)	(HK)	HU	IL	IT	JP	KR	NO	(PH)	PL	(SG)	SE	US
P1M18	AU	CAFR	CAEN	EN	FI	(HK)	(HU)	IL	IT	JP	(KR)	NO	(PH)	(PL)	(SG)	SE	US
P1M20	AU	(CAFR)	CAEN	(EN)	FI	(HK)	HU	(IL)	(IT)	(JP)	(KR)	(NO)	(PH)	PL	SG	SE	US
P1M21	AU	CAFR	CAEN	(EN)	FI	HK	HU	IL	(IT)	JP	KR	NO	PH	(PL)	(SG)	SE	US
P1M22	(AU)	(CAFR)	(CAEN)	(EN)	FI	HK	HU	IL	(IT)	(JP)	(KR)	(NO)	(PH)	(PL)	(SG)	SE	(US)
P1M23	(AU)	CAFR	CAEN	(EN)	FI	HK	HU	IL	(IT)	JP	KR	(NO)	PH	(PL)	(SG)	SE	US
P1M24	AU	CAFR	(CAEN)	(EN)	FI	(HK)	(HU)	IL	IT	(JP)	KR	NO	PH	(PL)	(SG)	SE	US

SISS = Second International Science Study, AU = Australia, CAEN = Canada English, CAFR = Canada French, EN = England, FI = Finland, HK = Hong, HU = Hungary, IL = Israel, IT = Italy, JP = Japan, KR = South Korea, NO = Norway, PH = Philippines, PL = Poland, SG = Singapore, SE = Sweden, US = United States

Table S14*Root Mean Square Difference (RMSD) Values of the Items in the First International Science Study*

Item	Thailand	Belgium (Flemish)	Belgium (French)	England	FR. Germany	Finland	Hungary	Italy	Japan	Netherlands	Scotland	Sweden	United States	n ^a
E1SA01	0.18	0.06	0.08	0.06	0.11	0.08	0.10	0.07	0.07	0.11	0.07	0.03	0.12	8
E1SA02	0.25	0.10	0.09	0.05	0.03	0.13	0.04	0.11	0.13	0.12	0.05	0.02	0.04	7
E1SA04	0.15	0.06	0.12	0.02	0.09	0.04	0.07	0.02	0.13	0.04	0.02	0.04	0.04	10
E1SA06	0.09	0.10	0.14	0.01	0.07	0.08	0.06	0.07	0.18	0.03	0.04	0.09	0.15	9
E1SA07	0.20	0.02	0.14	0.03	0.02	0.05	0.08	0.16	0.12	0.15	0.04	0.12	0.07	7
E1SA08	0.09	0.09	0.10	0.10	0.09	0.04	0.04	0.05	0.05	0.11	0.09	0.05	0.05	11
E1SA09	0.13	0.07	0.04	0.01	0.02	0.06	0.04	0.20	0.11	0.07	0.02	0.06	0.04	10
E1SA10	0.13	0.22	0.08	0.05	0.08	0.13	0.10	0.09	0.02	0.11	0.04	0.10	0.14	6
E1SA11	0.45	0.13	0.03	0.05	0.17	0.15	0.14	0.05	0.08	0.15	0.12	0.13	0.07	5
E1SA14	0.02	0.10	0.05	0.01	0.06	0.12	0.07	0.03	0.03	0.04	0.02	0.08	0.05	12
E1SA15	0.10	0.03	0.03	0.02	0.04	0.03	0.15	0.04	0.15	0.04	0.01	0.20	0.08	9
E1SA16	0.02	0.03	0.14	0.04	0.12	0.03	0.05	0.02	0.12	0.03	0.07	0.09	0.07	10
E1SA17	0.04	0.21	0.05	0.02	0.12	0.06	0.08	0.13	0.03	0.05	0.03	0.02	0.01	10
E1SA18	0.21	0.19	0.08	0.03	0.10	0.14	0.03	0.08	0.14	0.08	0.03	0.03	0.08	8
E1SA19	0.08	0.07	0.07	0.03	0.03	0.09	0.17	0.03	0.11	0.03	0.03	0.04	0.06	11
E1SA20	0.02	0.06	0.08	0.07	0.02	0.04	0.04	0.15	0.07	0.03	0.03	0.11	0.09	11

Note. Items with RMSD values higher than 0.1 are highlighted in bold.

^a Number of educational systems where the item exhibited a good fit (RMSD values less than 0.1)

Table S14*Root Mean Square Difference (RMSD) Values of the Items in the First International Science Study (cont.)*

Item	Thailand	Belgium (Flemish)	Belgium (French)	England	FR. Germany	Finland	Hungary	Italy	Japan	Netherlands	Scotland	Sweden	United States	n ^a
E1SB01	0.08	0.09	0.04	0.10	0.04	0.04	0.11	0.10	0.02	0.15	0.08	0.05	0.03	11
E1SB02	0.10	0.06	0.03	0.04	0.03	0.04	0.03	0.07	0.15	0.08	0.07	0.02	0.02	11
E1SB03	0.08	0.08	0.12	0.02	0.08	0.06	0.11	0.05	0.04	0.06	0.06	0.05	0.07	11
E1SB04	0.08	0.07	0.12	0.02	0.22	0.11	0.03	0.06	0.04	0.10	0.02	0.03	0.09	10
E1SB05	0.03	0.05	0.03	0.04	0.12	0.06	0.06	0.08	0.08	0.03	0.05	0.03	0.04	12
E1SB06	0.03	0.10	0.11	0.13	0.07	0.12	0.05	0.12	0.05	0.06	0.12	0.05	0.11	7
E1SB07	0.09	0.05	0.03	0.08	0.05	0.06	0.14	0.08	0.03	0.05	0.04	0.03	0.09	12
E1SB10	0.16	0.03	0.02	0.05	0.04	0.04	0.14	0.04	0.14	0.03	0.05	0.06	0.03	10
E1SB11	0.23	0.16	0.14	0.05	0.06	0.03	0.17	0.02	0.09	0.08	0.07	0.07	0.16	8
E1SB12	0.04	0.07	0.04	0.03	0.06	0.03	0.04	0.07	0.18	0.03	0.05	0.06	0.05	12
E1SB13	0.20	0.07	0.05	0.04	0.08	0.05	0.06	0.09	0.18	0.10	0.02	0.07	0.07	10
E1SB14	0.27	0.09	0.07	0.03	0.12	0.08	0.10	0.05	0.30	0.06	0.05	0.08	0.12	8
E1SB15	0.07	0.02	0.11	0.03	0.05	0.07	0.04	0.08	0.04	0.04	0.02	0.06	0.04	12
E1SB16	0.16	0.03	0.05	0.06	0.08	0.12	0.17	0.08	0.04	0.12	0.04	0.26	0.03	8
E1SB17	0.13	0.11	0.11	0.09	0.13	0.03	0.08	0.09	0.07	0.12	0.04	0.12	0.20	6
E1SB18	0.07	0.15	0.03	0.08	0.08	0.09	0.08	0.16	0.22	0.03	0.09	0.04	0.08	10
E1SB19	0.10	0.09	0.11	0.10	0.10	0.09	0.02	0.05	0.08	0.08	0.03	0.08	0.07	10
E1SB20	0.12	0.03	0.05	0.10	0.05	0.08	0.09	0.02	0.23	0.05	0.14	0.07	0.02	10

Note. Items with RMSD values higher than 0.1 are highlighted in bold.

^a Number of educational systems where the item exhibited a good fit (RMSD values less than 0.1)

Table S15

Root Mean Square Difference (RMSD) Values of the Items in the Second International Science Study

Item	Philippines	Australia	Canada (Fr)	Canada (Eng)	England	Finland	Hong Kong	Hungary	Israel	Italy	Japan	South Korea	Norway	Poland	Singapore	Sweden ^a	United States ^b	n ^c
P1M01	0.15	0.08	0.17	0.04	0.09	0.05	0.10	0.05	0.13	0.06	0.03	0.07	0.07	0.15	0.06	0.06	0.08	13
P1M02	0.05	0.10	0.03	0.05	0.14	0.02	0.03	0.04	0.06	0.14	0.12	0.01	0.06	0.08	0.01	0.06	0.03	13
P1M03	0.07	0.03	0.06	0.05	0.03	0.08	0.02	0.03	0.10	0.02	0.04	0.07	0.05	0.05	0.06	0.04	0.07	16
P1M04	0.05	0.02	0.02	0.01	0.04	0.02	0.04	0.01	0.06	0.03	0.05	0.01	0.03	0.08	0.07	0.03	0.04	17
P1M06	0.05	0.04	0.00	0.10	0.03	0.12	0.02	0.07	0.18	0.02	0.17	0.05	0.10	0.02	0.05	0.03	0.15	11
P1M07	0.15	0.10	0.03	0.04	0.05	0.04	0.11	0.27	0.13	0.10	0.05	0.14	0.08	0.34	0.09	0.07	0.04	10
P1M08	0.09	0.01	0.07	0.03	0.02	0.10	0.05	0.03	0.17	0.09	0.07	0.01	0.07	0.20	0.10	0.02	0.01	13
P1M09	0.04	0.02	0.09	0.04	0.02	0.02	0.07	0.05	0.04	0.06	0.06	0.12	0.05	0.08	0.08	0.05	0.04	16
P1M10	0.08	0.06	0.05	0.05	0.09	0.11	0.09	0.05	0.07	0.10	0.01	0.06	0.23	0.07	0.11	0.21	0.05	13
P1M11	0.08	0.03	0.04	0.16	0.02	0.03	0.02	0.18	0.20	0.03	0.04	0.06	0.11	0.05	0.18	0.12	0.08	11
P1M12	0.08	0.09	0.13	0.15	0.16	0.05	0.01	0.09	0.17	0.04	0.08	0.18	0.21	0.10	0.03	0.08	0.05	11
P1M13	0.09	0.03	0.04	0.03	0.06	0.14	0.09	0.24	0.09	0.04	0.11	0.04	0.34	0.01	0.05	0.14	0.03	12
P1M14	0.05	0.02	0.02	0.12	0.04	0.05	0.07	0.03	0.09	0.08	0.08	0.03	0.16	0.05	0.04	0.06	0.17	14
P1M15	0.02	0.06	0.03	0.02	0.05	0.05	0.03	0.12	0.06	0.12	0.10	0.07	0.17	0.14	0.09	0.12	0.01	11
P1M16	0.03	0.05	0.04	0.05	0.09	0.10	0.06	0.07	0.09	0.09	0.16	0.01	0.08	0.11	0.09	0.05	0.04	15
P1M17	0.05	0.02	0.04	0.05	0.02	0.05	0.07	0.05	0.04	0.05	0.15	0.02	0.05	0.04	0.04	0.01	0.15	15
P1M18	0.05	0.08	0.07	0.07	0.10	0.06	0.06	0.09	0.10	0.07	0.05	0.10	0.11	0.15	0.01	0.06	0.09	14
P1M20	0.11	0.02	0.08	0.06	0.10	0.02	0.07	0.05	0.23	0.10	0.04	0.10	0.05	0.06	0.04	0.01	0.05	14
P1M21	0.03	0.08	0.08	0.06	0.12	0.07	0.20	0.01	0.03	0.08	0.03	0.05	0.14	0.07	0.09	0.08	0.06	14
P1M22	0.09	0.10	0.12	0.10	0.07	0.04	0.05	0.04	0.11	0.14	0.19	0.08	0.06	0.18	0.07	0.06	0.15	11
P1M23	0.03	0.15	0.09	0.11	0.12	0.01	0.02	0.03	0.03	0.21	0.01	0.06	0.16	0.12	0.10	0.05	0.14	9
P1M24	0.06	0.07	0.05	0.07	0.14	0.11	0.04	0.18	0.06	0.07	0.10	0.06	0.08	0.08	0.18	0.10	0.08	11

Note. Items with RMSD values higher than 0.1 are highlighted in bold.

^a Grade 4 in SISS

^b Phase 1 in SISS

^c Number of educational systems where the item exhibited a good fit (RMSD values less than 0.1)

Table S16

Ranking Positions of Educational Systems in FISS: Comparisons Between the Technical Report, Alignment-Derived Model, Restricted Model, and Mean Scores

Educational System ^a	Technical Report ^b	Alignment	Restricted	Mean Scores
Japan	1	1	1	1
Sweden	2	5	3	3
Belgium (Flemish)	3	2	2	2
United States	4	3	4	4
Finland	5	6	6	6
Hungary	6	9	8	8
Italy	7	4	5	5
England	8	7	7	9
Netherlands	9	11	10	10
FR. Germany	10	8	9	7
Scotland	11	10	12	12
Belgium (French)	12	12	11	11

FISS = First International Science Study.

^a There were originally two groups of educational systems. In our analysis, only Thailand remains in the second group, so that group is not shown in the table.

^b Retrieved from Table 7.2 in Comber and Keeves (1973).

Table S17

Ranking Positions of Educational Systems in SISS: Comparisons Between the Technical Report, Alignment-Derived Model, Restricted Model, and Mean Scores

Educational System	Technical Report ^a	Alignment	Restricted	Mean Scores
Japan	1	5	1	1
South Korea	2	1	4	3
Finland	3	4	2	2
Sweden ^b	4	2	3	4
Canada (French)	5	8	6	7
Hungary	6	3	5	5
Canada (English)	7	9	9	9
Italy	8	6	8	8
Australia	10	11	11	11
Norway	11	10	12	12
Poland	12	13	13	13
Israel	13	12	10	10
England	14	15	14	14
Singapore	15	16	16	16
Hong Kong	16	14	15	15
Philippines	17	17	17	17
Nigeria	18	–	–	–

Note. “–” indicates that the educational system was removed from the analysis for having at least one item with a negative point biserial correlation.

The United States is not shown as we used only phase 1 data, whereas the original SISS study used both phase 1 and phase 2 data.

SISS = Second International Science Study.

^a Retrieved from Table 3.1 in Postlethwaite and Wiley (1992)

^b Grade 4 in SISS.

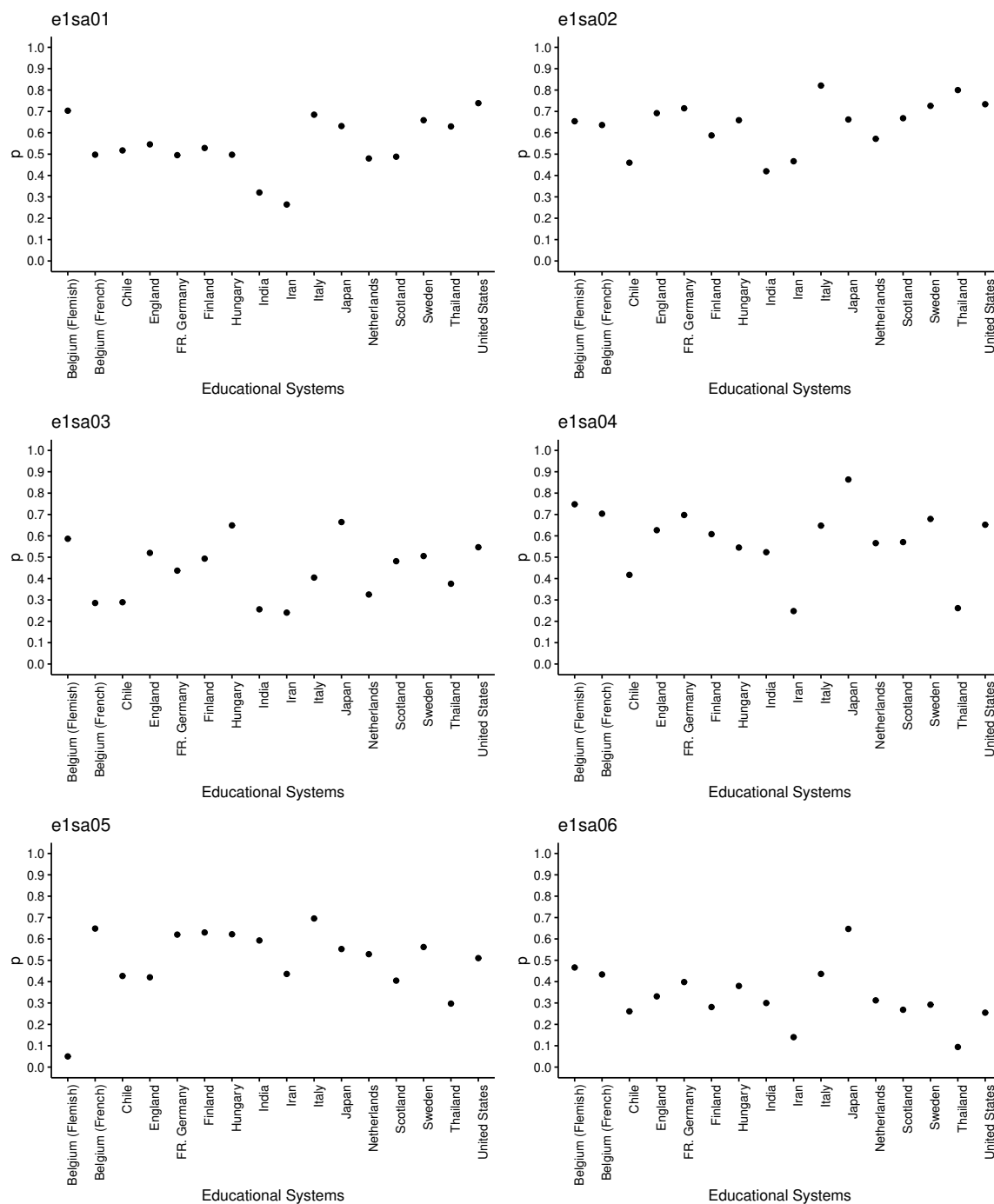
Figure S1*Difficulty Levels of the First International Science Study Items*

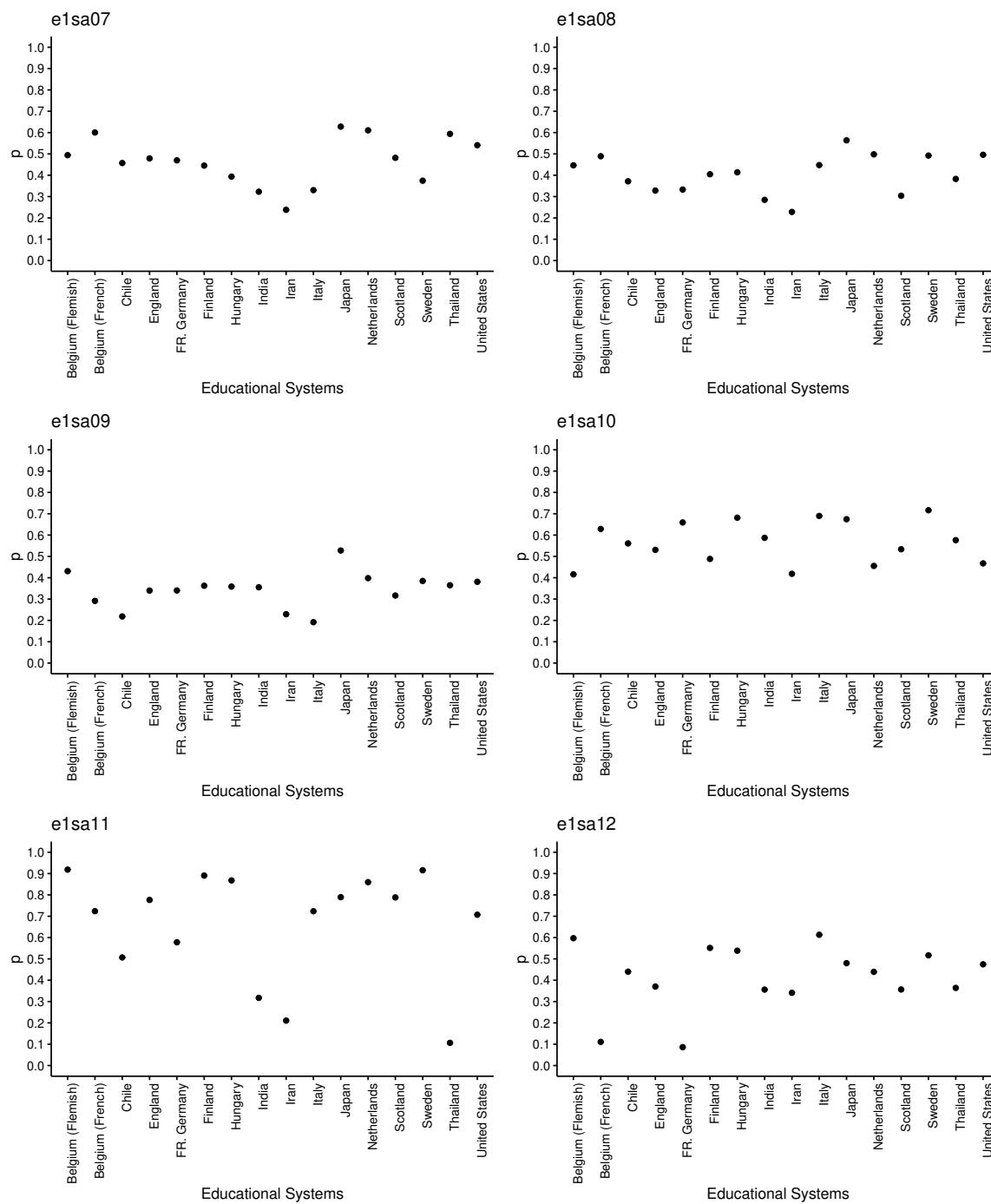
Figure S1*Difficulty Levels of the First International Science Study Items (cont.)*

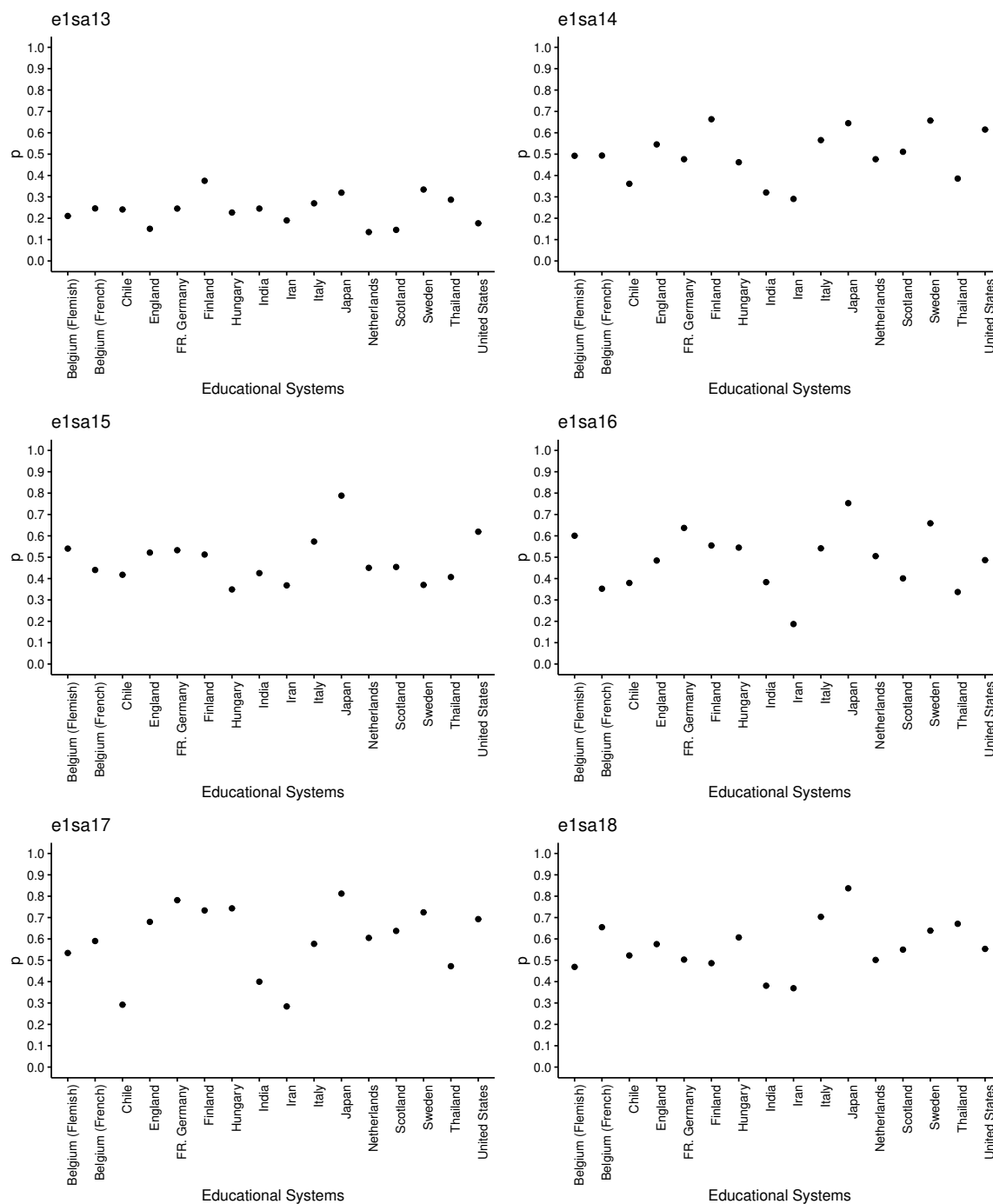
Figure S1*Difficulty Levels of the First International Science Study Items (cont.)*

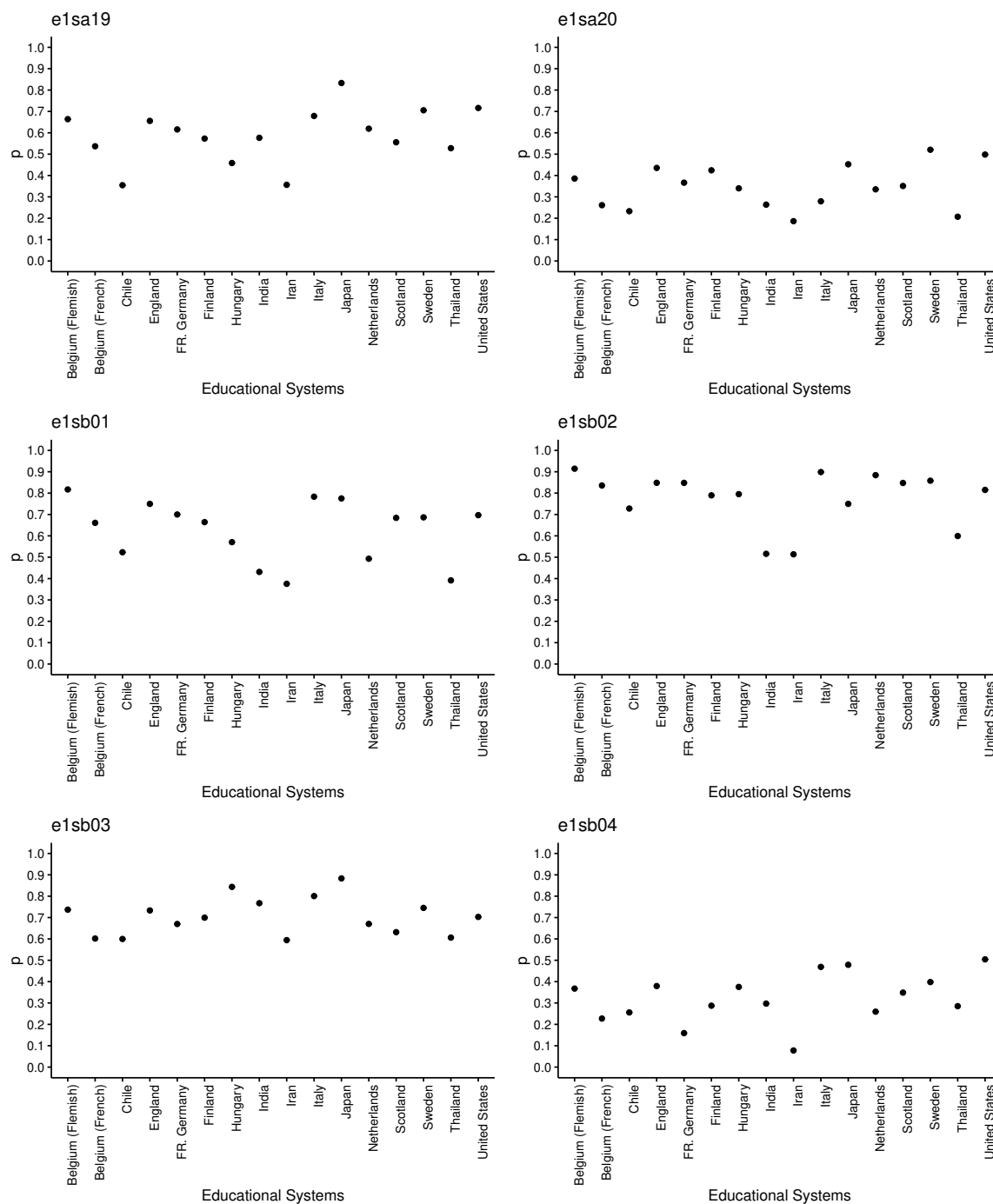
Figure S1*Difficulty Levels of the First International Science Study Items (cont.)*

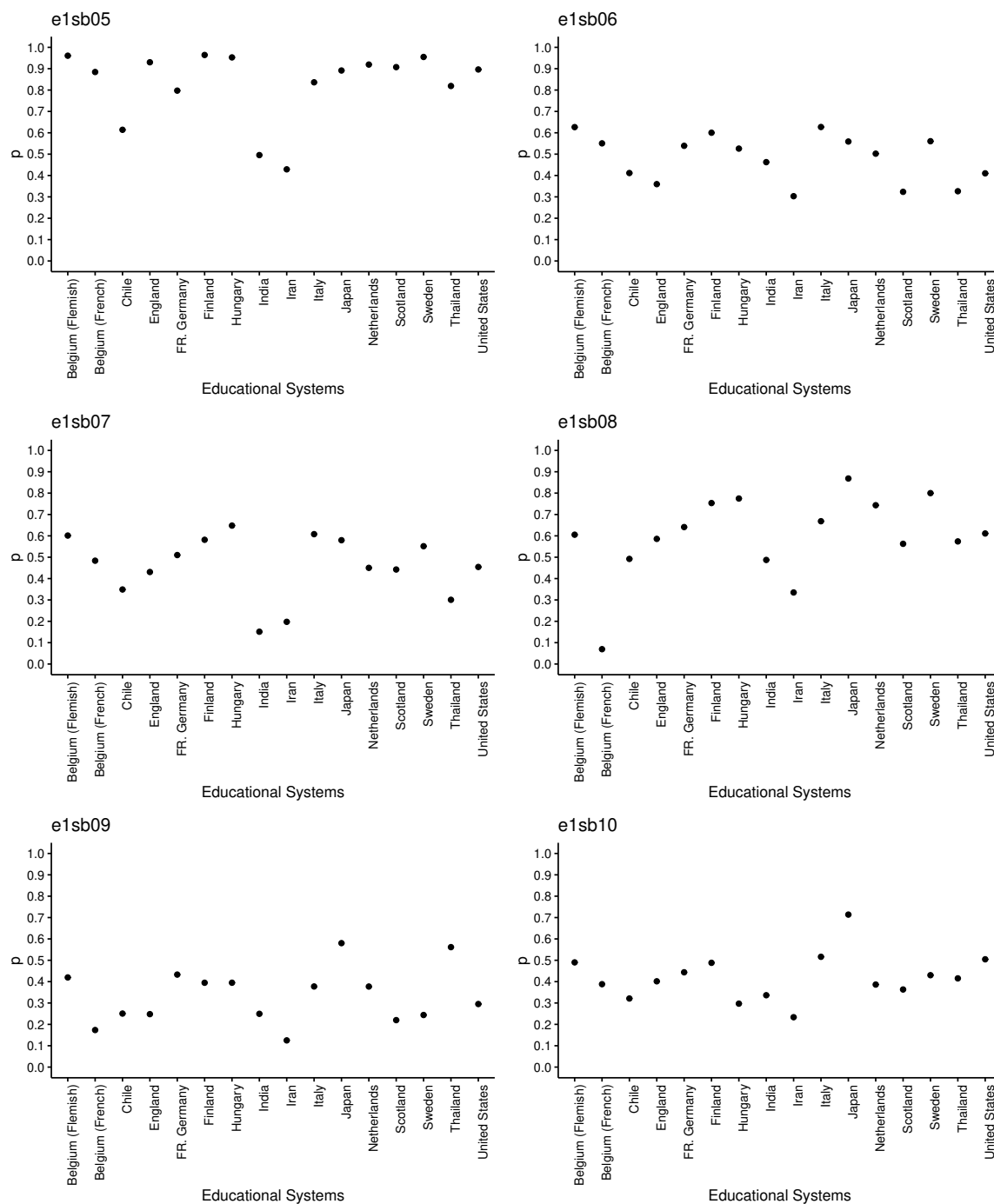
Figure S1*Difficulty Levels of the First International Science Study Items (cont.)*

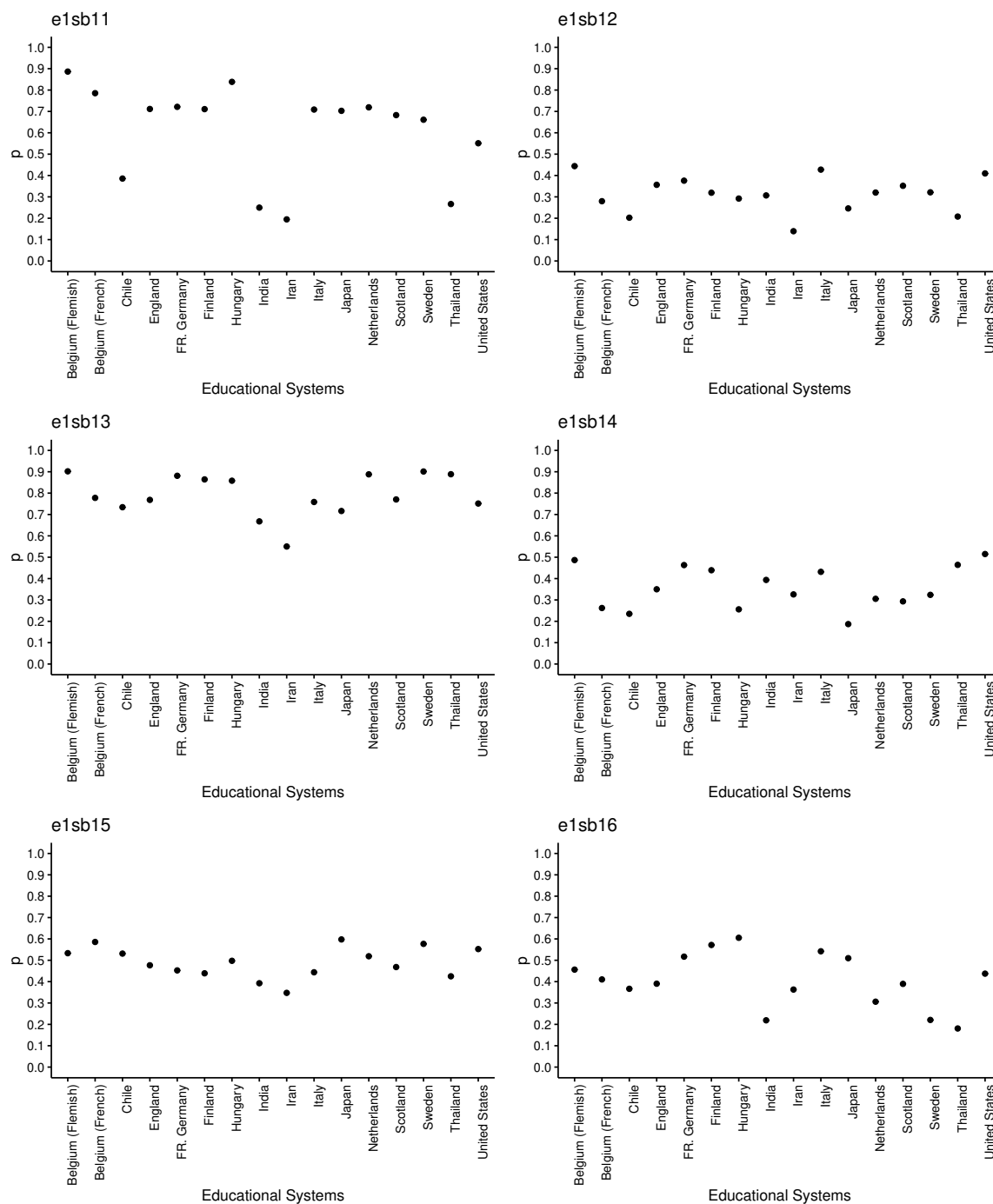
Figure S1*Difficulty Levels of the First International Science Study Items (cont.)*

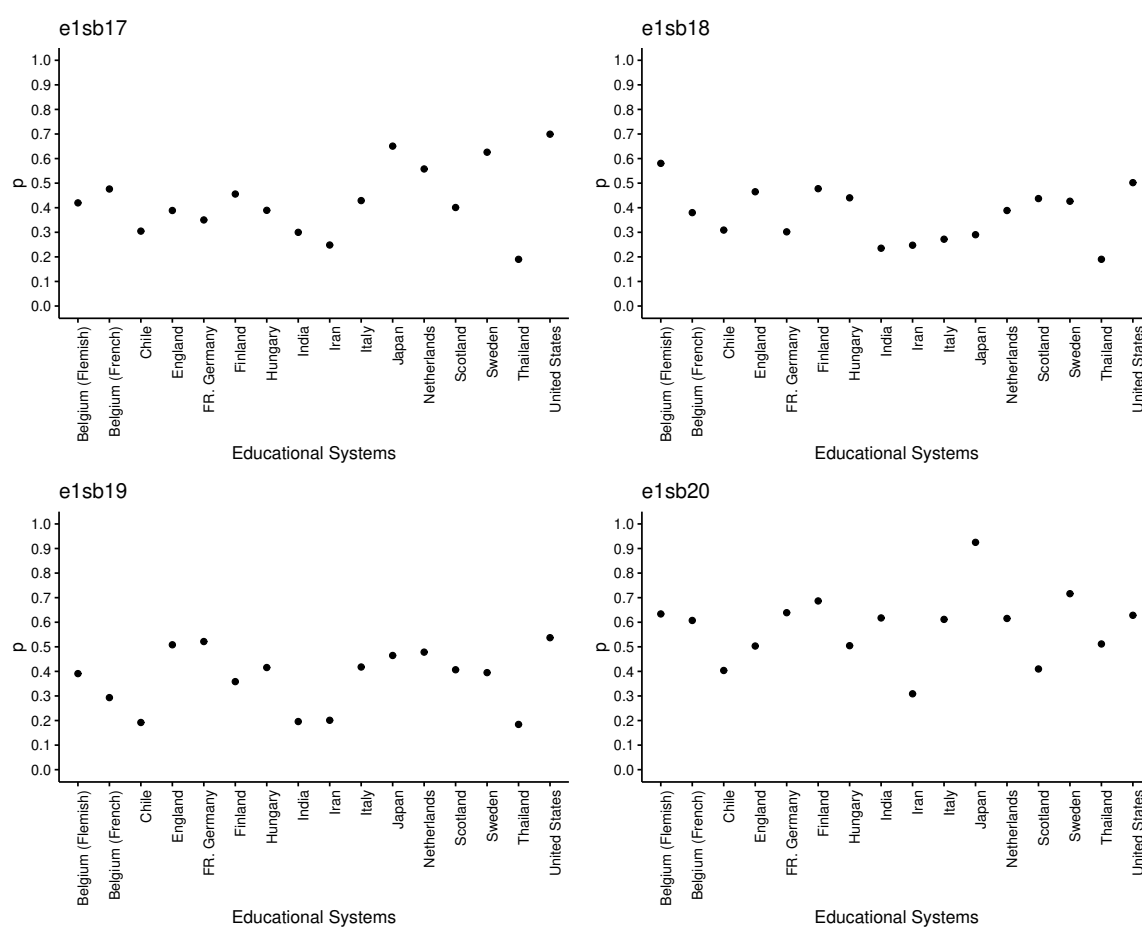
Figure S1*Difficulty Levels of the First International Science Study Items (cont.)*

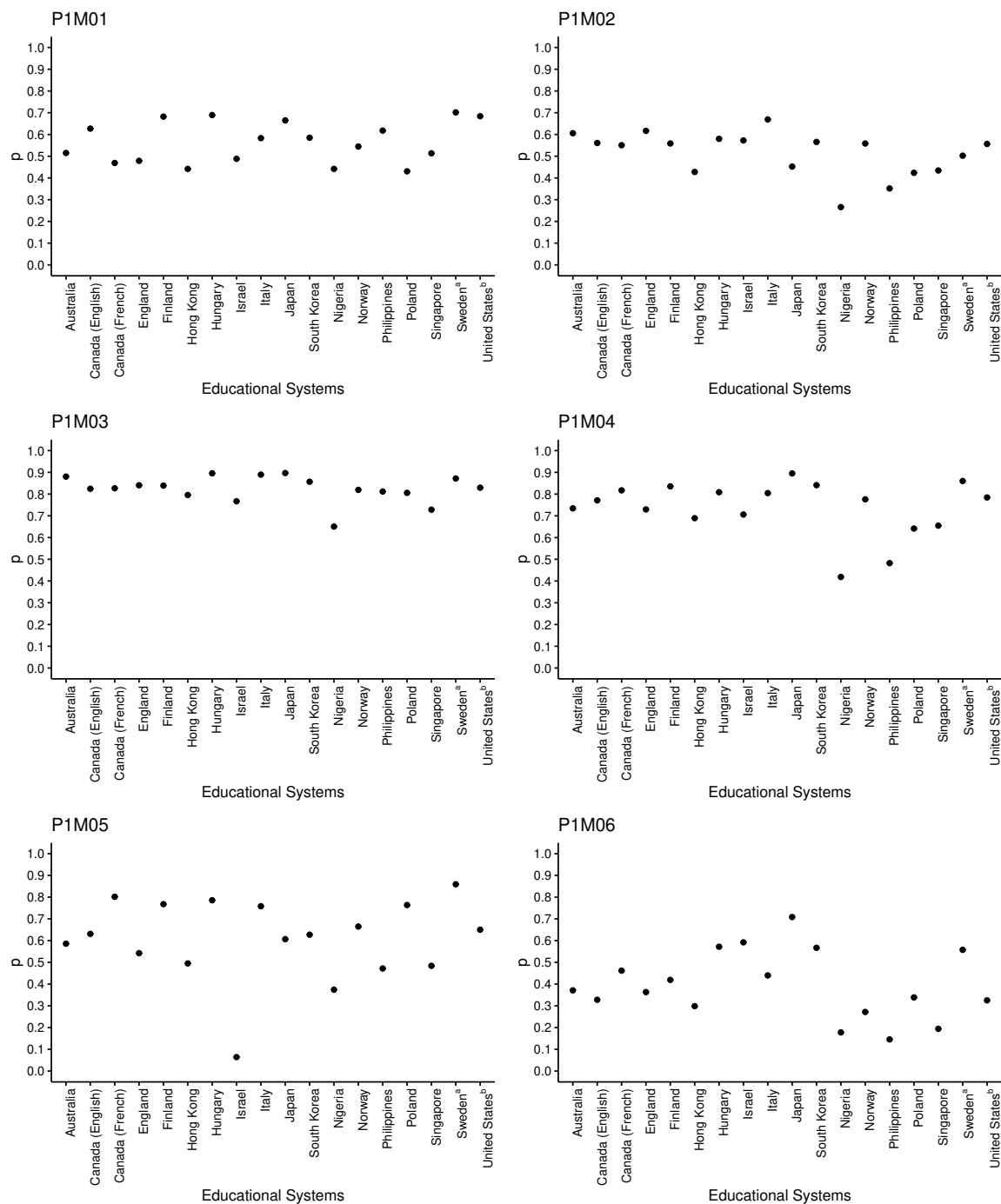
Figure S2*Difficulty Levels of the Second International Science Study Items*

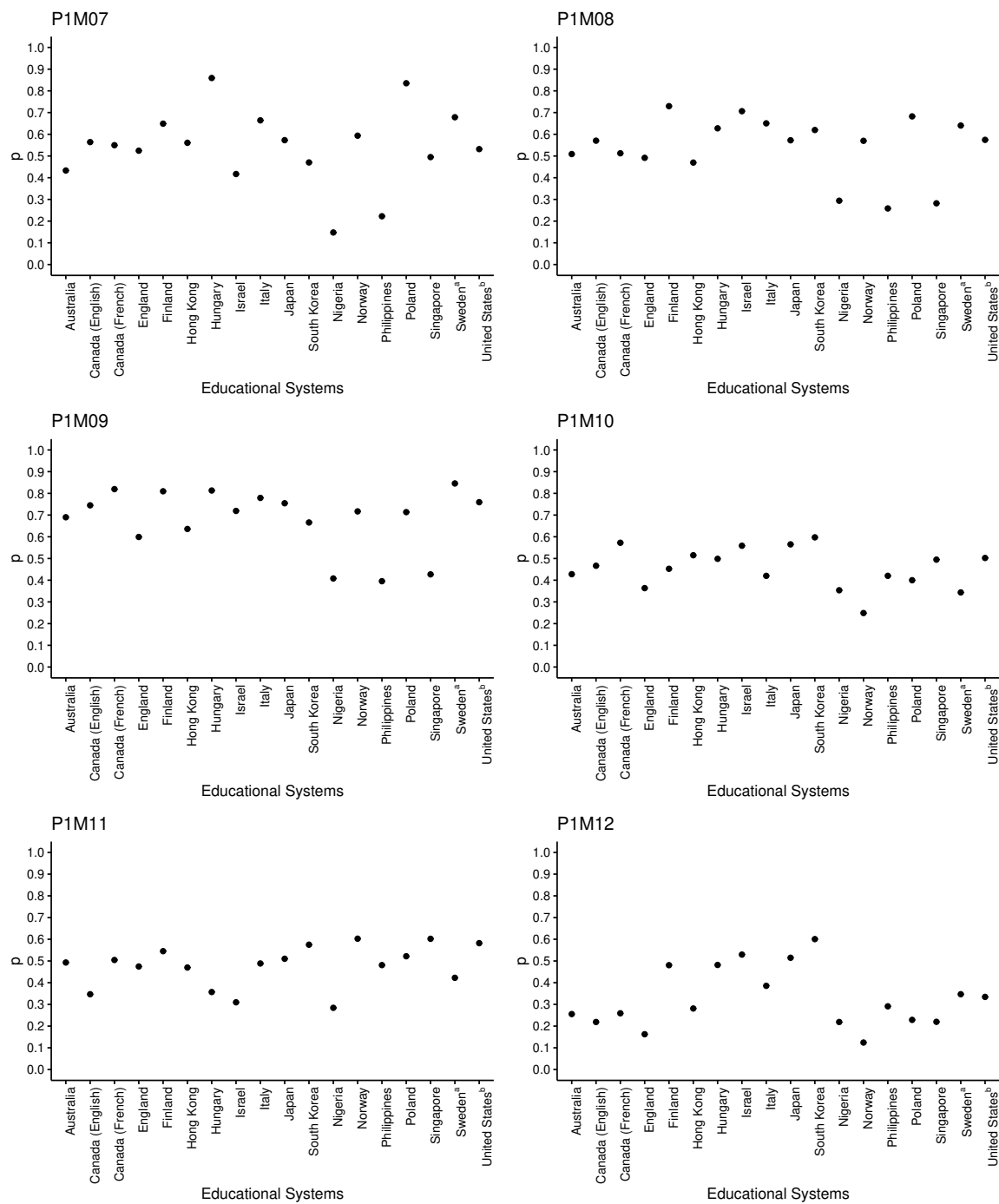
Figure S2*Difficulty Levels of the Second International Science Study Items (cont.)*

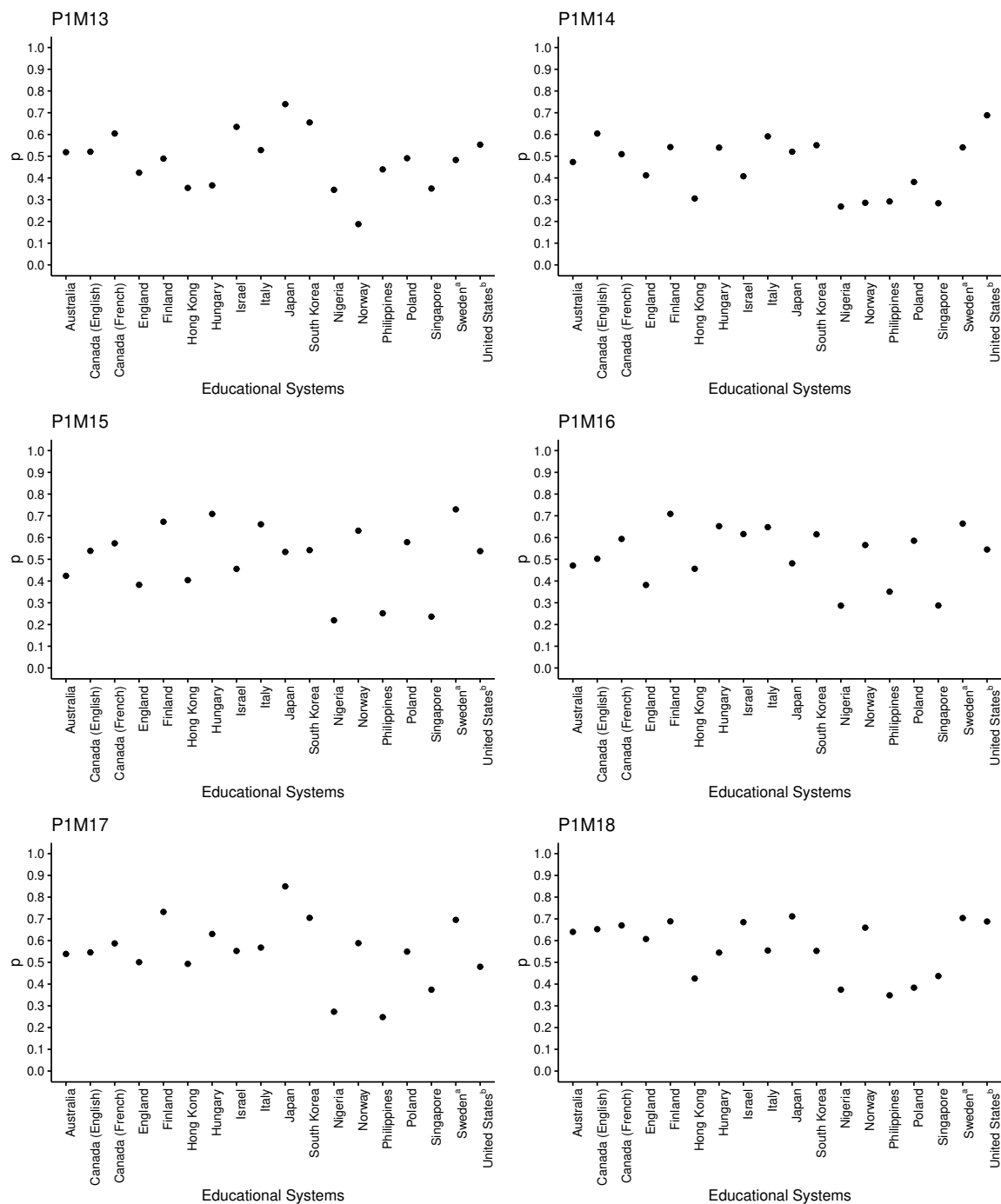
Figure S2*Difficulty Levels of the Second International Science Study Items (cont.)*

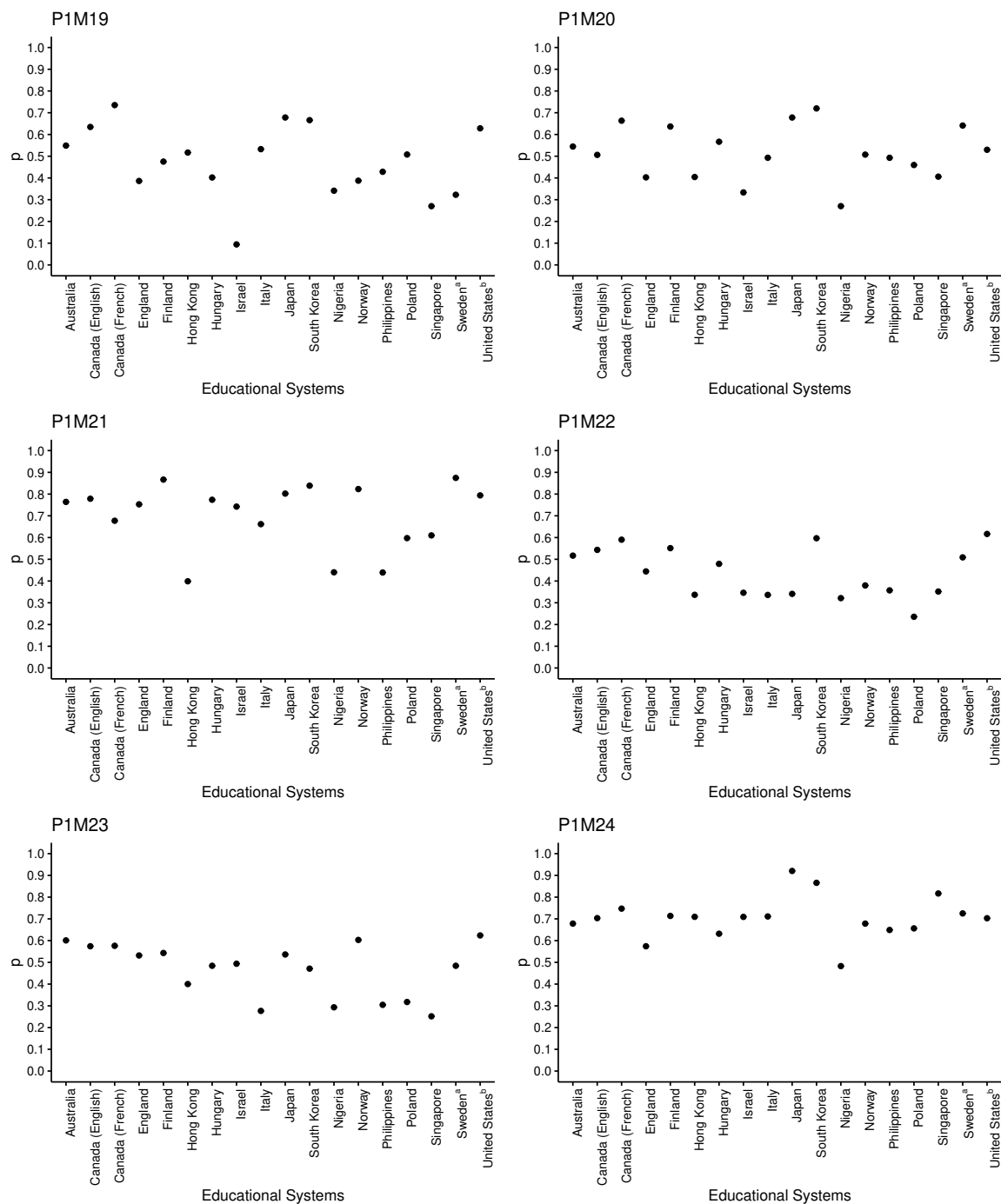
Figure S2*Difficulty Levels of the Second International Science Study Items (cont.)*

Figure S3

Point Biserial Correlations map in the First International Science Study

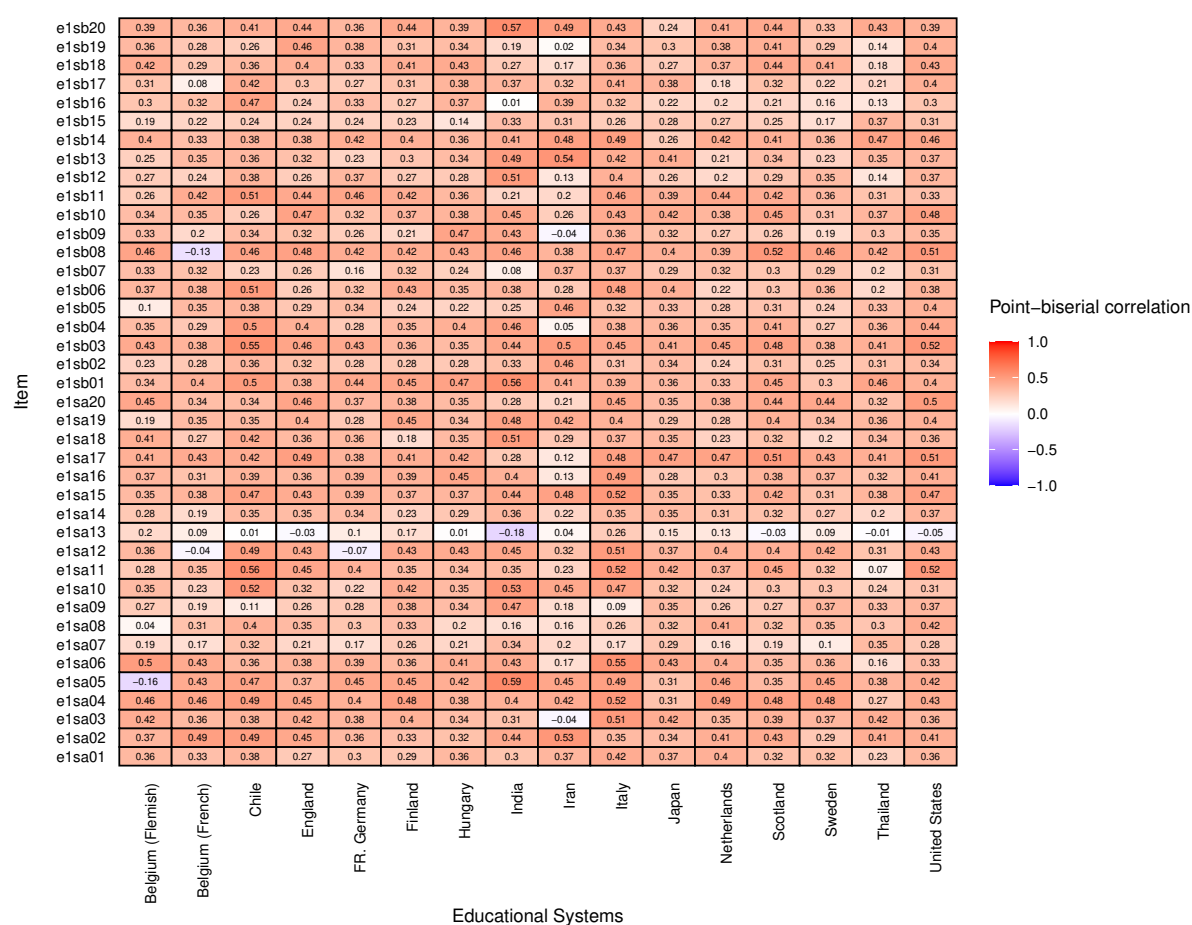


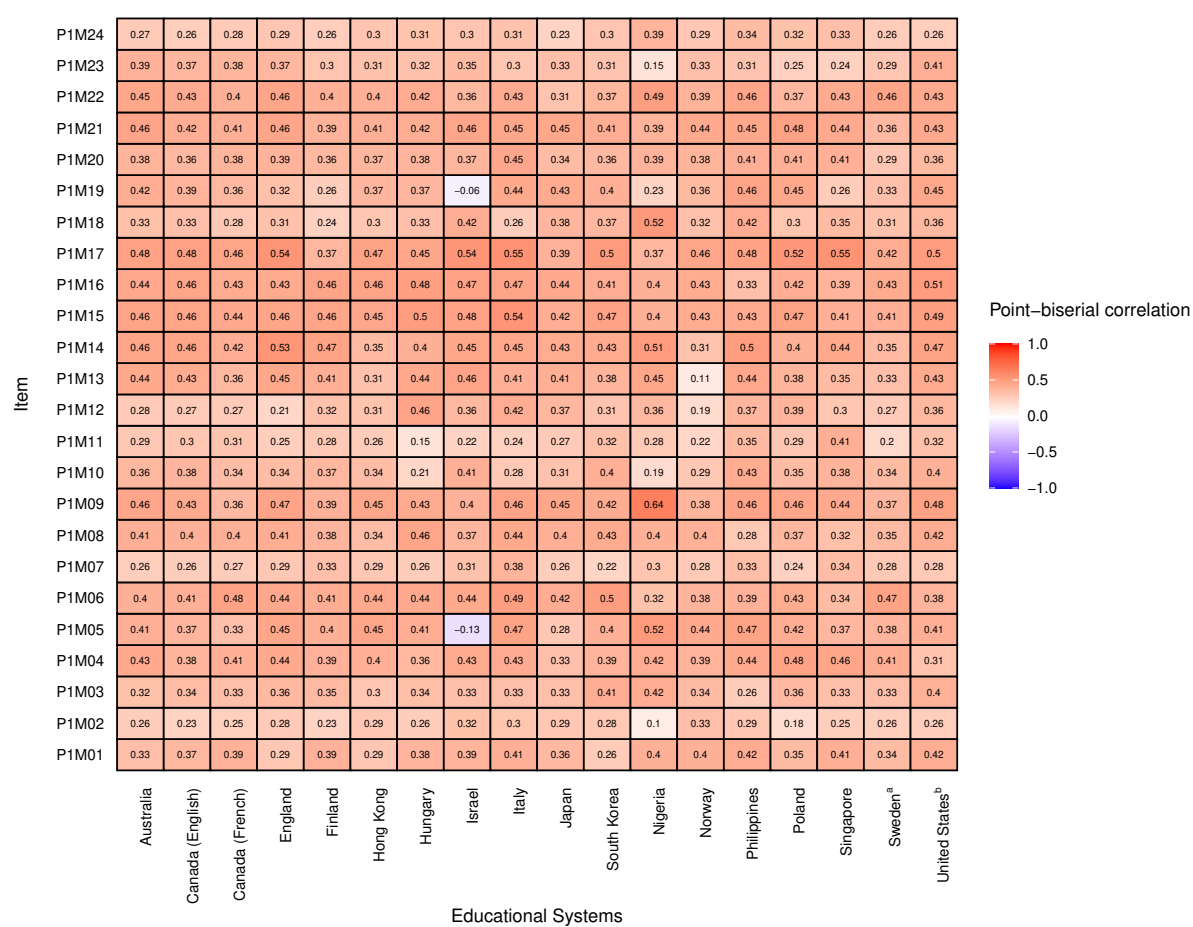
Figure S4*Point Biserial Correlations map in the Second International Science Study*

Figure S5

Comparison of Test Information Function in Each System Using Alignment-Derived Model and Restricted Model in the First International Science Study

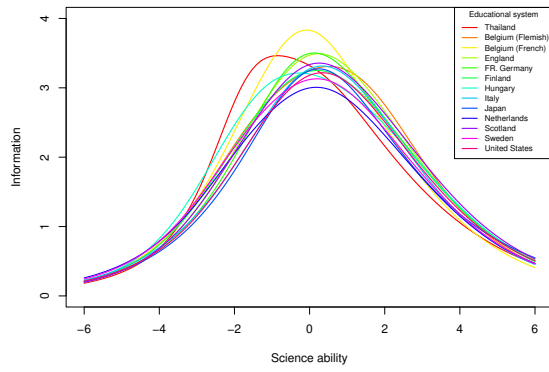
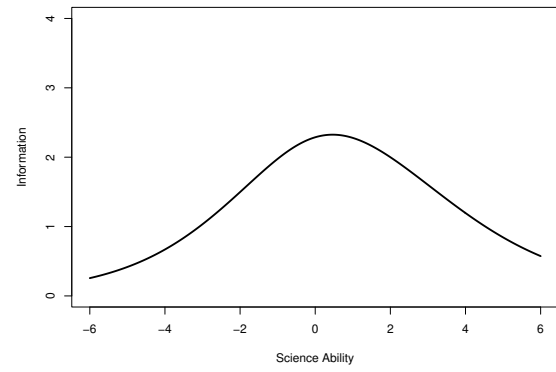
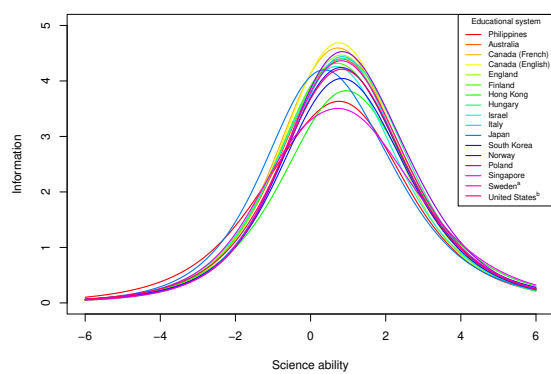
(a) *Alignment-Derived Model*(b) *Restricted Model*

Figure S6

Comparison of Test Information Function in Each System Using Alignment-Derived Model and Restricted Model in the Second International Science Study

(a) *Alignment-Derived Model*(b) *Restricted Model*