



Scaling Outcomes

International characteristics of the item pool	208
▪ Test targeting.....	208
▪ Test reliability.....	208
▪ Domain inter-correlations	208
▪ Science scales.....	215
Scaling outcomes.....	216
▪ National item deletions.....	216
▪ International scaling.....	219
▪ Generating student scale scores.....	219
Test length analysis.....	219
Booklet effects.....	221
▪ Overview of the PISA cognitive reporting scales.....	232
▪ PISA overall literacy scales	234
▪ PISA literacy scales.....	234
▪ Special purpose scales.....	234
Observations concerning the construction of the PISA overall literacy scales.....	235
▪ Framework development.....	235
▪ Testing time and item characteristics	236
▪ Characteristics of each of the links	237
Transforming the plausible values to PISA scales.....	246
▪ Reading.....	246
▪ Mathematics.....	246
▪ Science.....	246
▪ Attitudinal scales	247
Link error.....	247



INTERNATIONAL CHARACTERISTICS OF THE ITEM POOL

When main study data were received from each participating country, they were first verified and cleaned using the procedures outlined in Chapter 10. Files containing the achievement data were prepared and national-level Rasch and traditional test analyses were undertaken. The results of these analyses were included in the reports that were returned to each participating country (see Chapter 9).

After processing at the national level, a set of international-level analyses was undertaken. Some involved summarising national analyses, while others required an analysis of the international data set.

The final international cognitive data set (that is, the data set of coded achievement booklet responses) consisted of 398 750 students from 57 participating countries. Table 12.1 shows the total number of sampled students, broken down by participating country and test booklet.

Test targeting

Each of the domains was separately scaled to examine the targeting of the tests. Figure 12.1 shows the match between the international (OECD countries only) item difficulty distribution and the distribution of OECD's student achievement for each of reading, mathematics, science, interest and support, respectively. The figures consist of two panels. The left panel, students, shows the distribution of students' Rasch-scaled achievement estimates. Students at the top end of this distribution have higher proficiency estimates than students at the lower end of the distribution. The right panel, item difficulties, shows the distribution of Rasch-estimated item difficulties.

In each of Figure 12.2 to Figure 12.5 the student proficiency distribution, shown by X_s , is well matched to the item difficulty distribution. For the interest scale (Figure 12.4) the items are well-matched in terms of average difficulty, but the items are not widely dispersed on the scale. For the support items, shown in Figure 12.5 shows it is clear that the items were very easy to agree with for students in OECD countries. The figures are constructed so that when a student and an item are located at the same height on the scale then the student has a 50% chance of responding correctly to the item.

Test reliability

A second test characteristic that is of importance is the test reliability. Table 12.2 shows the reliability for each of the five overall scales (mathematical literacy, reading literacy, combined scientific literacy and the attitude scales interest and support) before conditioning and based upon five separate scalings, using plausible values and using WLEs. The reliabilities for the minor domains are higher when using WLEs, because students that were not assessed in mathematics or reading were excluded from the calculation of the WLE reliabilities. These students do get plausible values, but there is no information available about these students (no scores on other domains, because of using uni-dimensional models and no background information because these are the reliabilities before conditioning). The international reliability for each domain after conditioning is reported later in Table 12.6.

Domain inter-correlations

Correlations between the ability estimates for individual students in each of the five domains, the latent correlations, as estimated by ConQuest® (Wu, Adams and Wilson, 1997) are given in Table 12.3. It is important to note that these latent correlations are unbiased estimates of the true correlation between the underlying latent variables. As such they are not attenuated by the unreliability of the measures and will generally be higher than the typical product moment correlations that have not been disattenuated for unreliability. The results in Table 12.3 are reported for both OECD countries and for all participating countries.¹



Table 12.1
Number of sampled student by country and booklet

	Booklet													UH	Total		
	1	2	3	4	5	6	7	8	9	10	11	12	13				
OECD	Australia	1088	1058	1081	1118	1124	1092	1074	1071	1066	1086	1106	1092	1114		14170	
	Austria	370	386	383	394	395	378	372	366	370	371	387	369	367	19	4927	
	Belgium	672	668	680	647	665	660	677	662	674	676	671	672	661	172	8857	
	Canada	1725	1738	1692	1736	1744	1768	1762	1769	1738	1768	1745	1717	1744		22646	
	Czech Republic	449	463	438	457	442	437	439	452	451	446	439	447	450	122	5932	
	Denmark	355	362	348	349	367	366	361	339	335	334	340	331	345		4532	
	Finland	361	372	366	366	358	349	361	359	362	367	367	366	360		4714	
	France	361	360	358	356	358	361	368	363	362	367	371	364	367		4716	
	Germany	366	361	360	363	354	361	364	366	370	368	371	360	367	160	4891	
	Greece	370	381	390	381	384	378	367	359	370	367	383	366	377		4873	
	Hungary	355	356	359	345	342	345	332	342	327	341	351	347	348		4490	
	Iceland	297	294	290	303	299	298	289	284	265	288	294	289	299		3789	
	Ireland	341	360	362	347	348	347	350	363	351	351	344	357	364		4585	
	Italy	1674	1666	1671	1686	1656	1661	1684	1712	1716	1678	1658	1665	1646		21773	
	Japan	469	471	467	451	450	453	457	454	458	446	451	461	464		5952	
	Korea	396	407	400	389	389	385	394	405	406	404	408	394	399		5176	
	Luxembourg	362	367	358	353	349	351	352	355	346	340	345	347	342		4567	
	Mexico	2328	2391	2406	2415	2369	2356	2423	2356	2373	2402	2373	2379	2400		30971	
	Netherlands	368	358	366	363	358	373	368	375	366	372	372	360	370	102	4871	
	New Zealand	394	368	371	372	379	368	361	363	360	371	380	365	371		4823	
	Norway	368	356	366	367	369	369	366	351	352	353	362	355	358		4692	
	Poland	425	421	422	433	429	434	441	445	440	413	409	414	421		5547	
	Portugal	378	388	387	381	391	397	382	395	405	386	413	397	409		5109	
	Slovak Republic	373	368	362	365	369	365	364	353	350	355	358	359	352	38	4731	
	Spain	1538	1534	1514	1489	1479	1507	1522	1513	1485	1501	1499	1505	1518		19604	
	Sweden	337	336	346	346	344	336	336	336	336	350	360	342	338		4443	
	Switzerland	932	953	933	931	940	932	951	939	948	925	927	950	931		12192	
	Turkey	380	381	379	377	377	379	376	377	376	382	391	381	386		4942	
	United Kingdom	993	996	997	997	1008	1034	1034	1020	1003	1027	1011	1021	1011		13152	
	United States	437	438	425	427	419	419	434	423	430	420	443	451	445		5611	
	Partners	Argentina	347	340	332	336	340	333	325	328	330	322	331	333	342		4339
		Azerbaijan	404	407	399	407	405	407	387	387	391	398	399	397	396		5184
		Brazil	702	723	701	737	713	718	719	704	714	693	739	711	721		9295
		Bulgaria	342	345	349	342	351	347	344	349	350	347	346	345	341		4498
Chile		413	391	395	395	414	412	410	392	394	401	401	408	407		5233	
Colombia		336	350	338	350	350	347	346	341	349	355	337	336	343		4478	
Croatia		414	421	412	410	407	404	398	399	389	392	381	378	408		5213	
Estonia		370	365	367	360	373	369	375	378	383	383	379	380	383		4865	
Hong Kong-China		348	349	353	349	357	362	371	360	363	368	368	352	345		4645	
Indonesia		822	812	806	812	816	804	815	820	825	837	825	833	820		10647	
Israel		362	357	346	363	364	350	345	351	341	335	346	352	372		4584	
Jordan		507	505	489	492	481	485	498	499	508	513	505	510	517		6509	
Kyrgyzstan		455	459	460	460	445	456	447	462	461	453	457	444	445		5904	
Latvia		368	357	359	353	359	364	362	356	369	354	368	380	370		4719	
Liechtenstein		27	26	25	26	26	27	26	25	26	25	26	26	28		339	
Lithuania		349	355	354	370	373	368	366	373	372	375	367	368	354		4744	
Macao-China		369	369	365	359	372	368	370	364	362	365	368	371	358		4760	
Montenegro		342	346	351	335	354	344	351	347	343	349	336	327	330		4455	
Qatar		476	471	488	483	478	493	477	475	482	480	496	493	473		6265	
Romania		392	394	401	406	399	396	385	388	393	393	394	388	389		5118	
Russian Federation		456	455	444	448	450	440	453	449	439	439	446	441	439		5799	
Serbia		369	370	372	363	371	375	371	364	379	377	358	365	364		4798	
Slovenia		487	508	488	505	497	485	491	492	480	488	483	484	490	217	6595	
Chinese Taipei		680	687	679	673	678	680	674	673	676	686	685	676	668		8815	
Thailand		474	475	468	466	469	483	488	478	473	474	482	486	476		6192	
Tunisia		360	343	361	368	364	361	363	364	342	349	349	356	360		4640	
Uruguay		387	390	389	367	374	368	368	348	369	368	366	364	381		4839	



Figure 12.1
Item plot for mathematics items

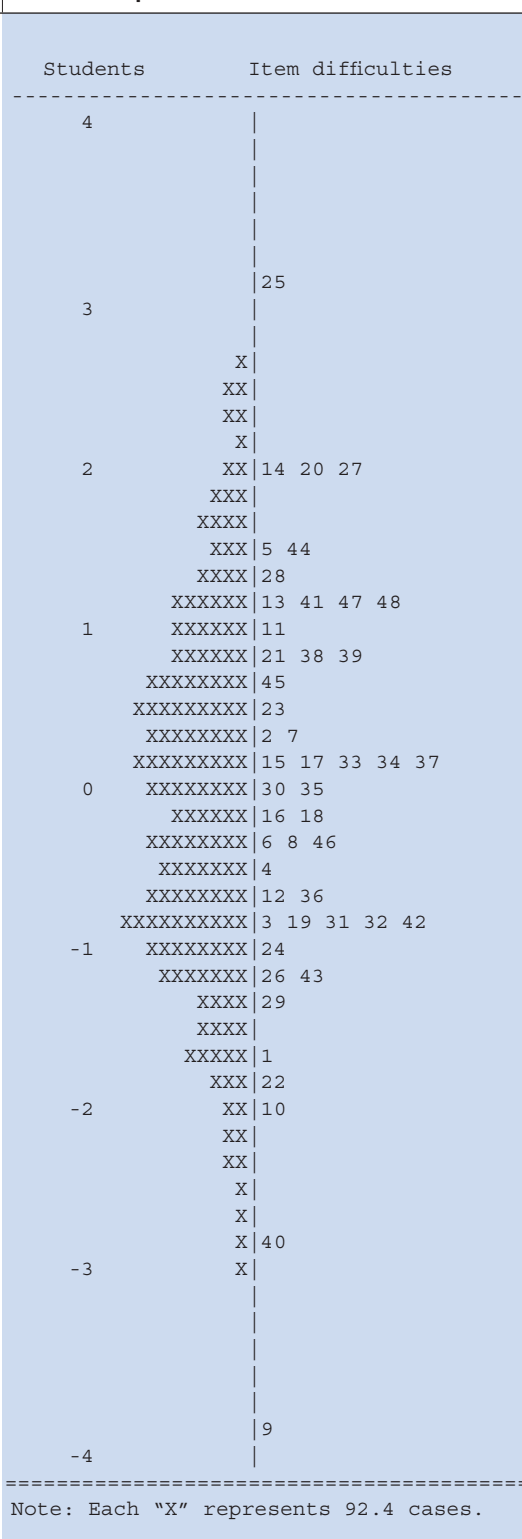




Figure 12.2
Item plot for reading items

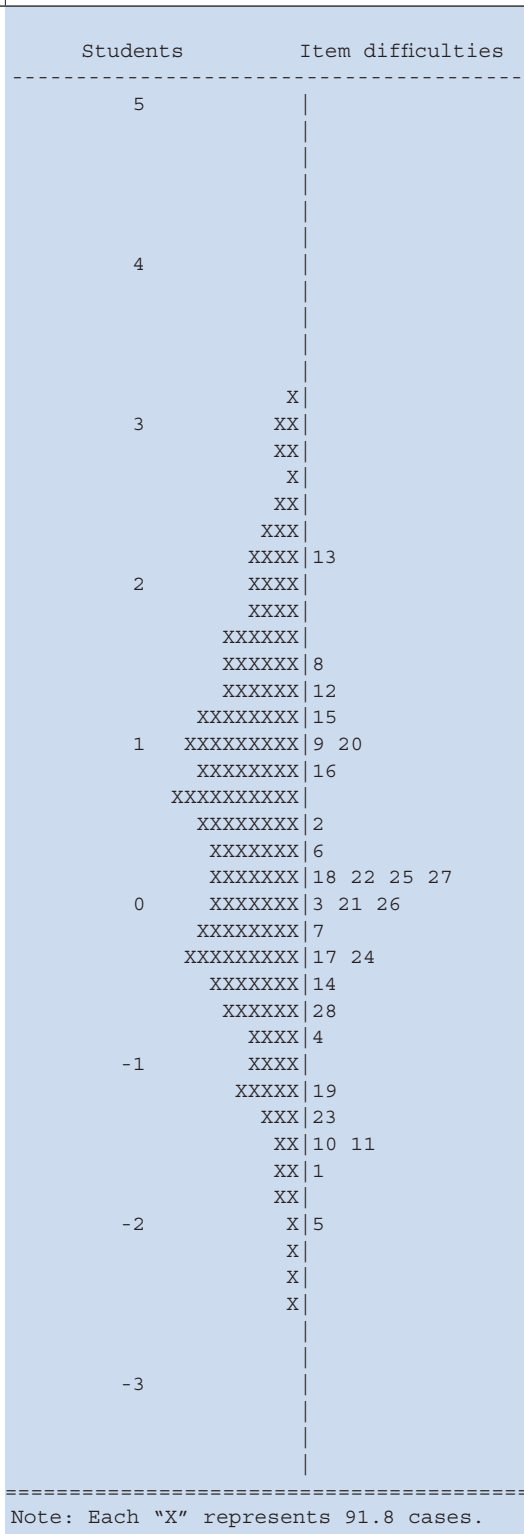




Figure 12.3
Item plot for science items

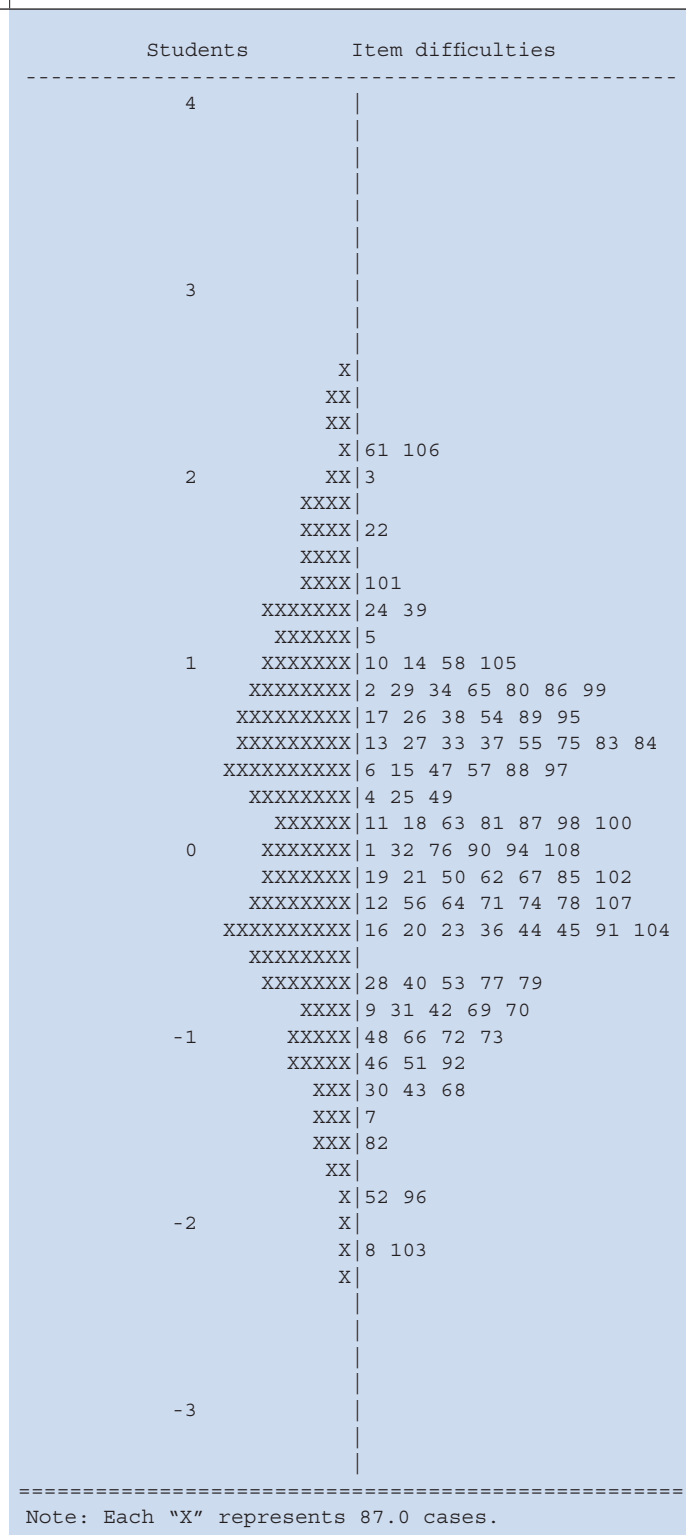




Figure 12.4
Item plot for interest items

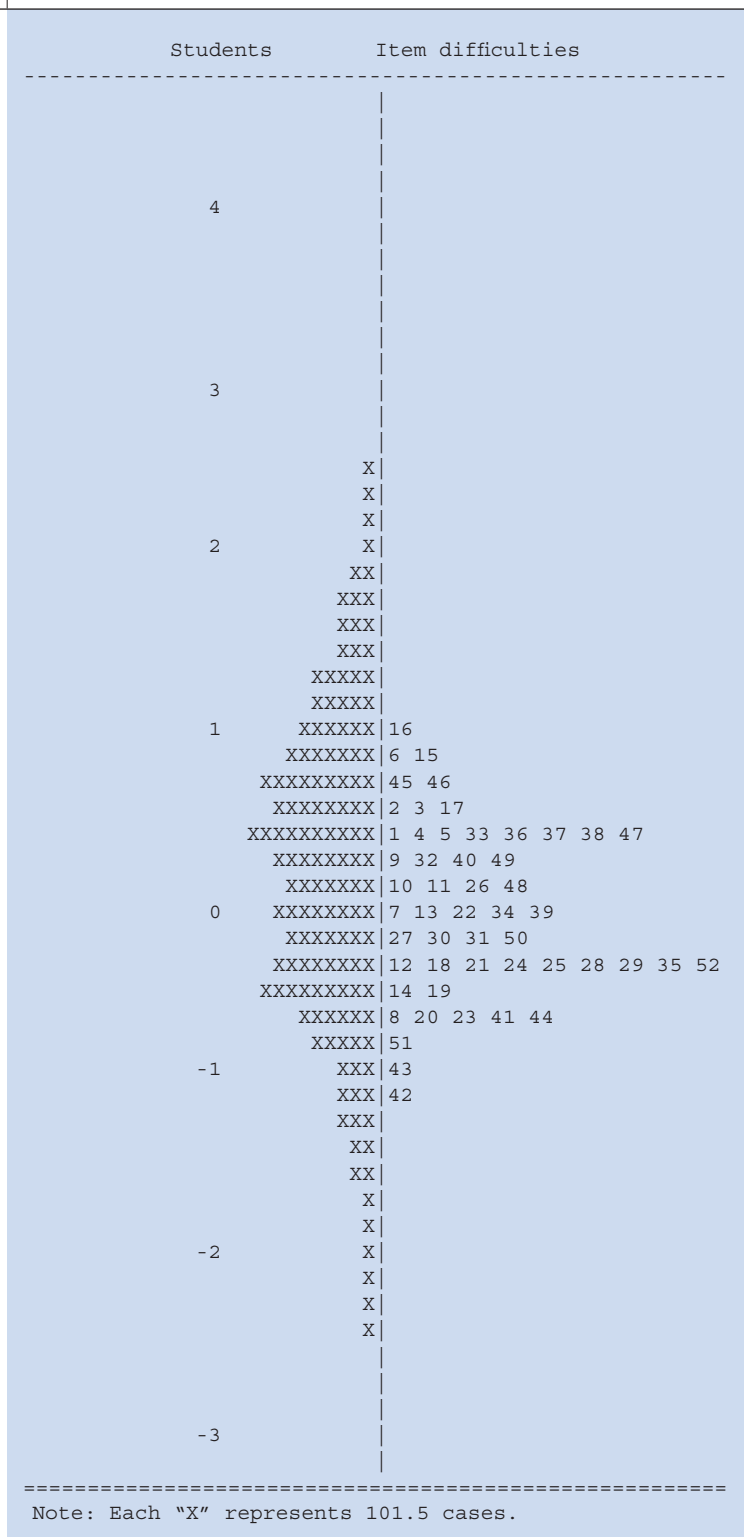




Figure 12.5
Item plot for support items

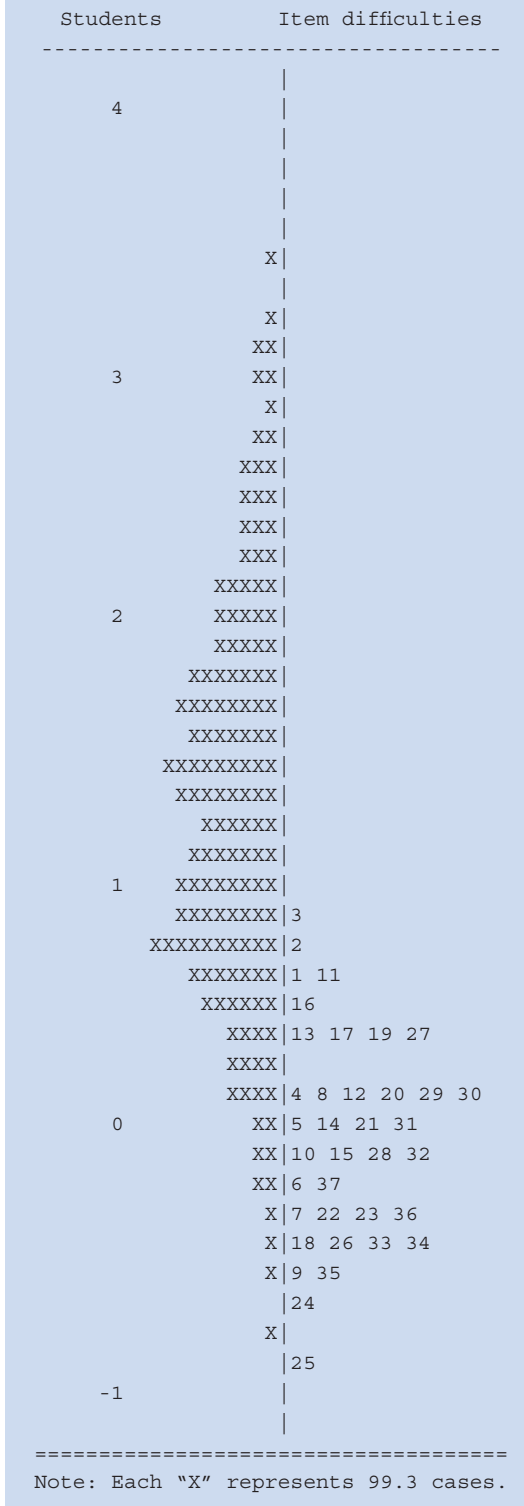




Table 12.2
Reliabilities of each of the four overall scales when scaled separately

Domain	Reliability (PV)	Reliability (WLE)
Mathematics	0.613	0.784
Reading	0.429	0.780
Science	0.856	0.832
Interest	0.886	0.867
Support	0.725	0.705

Table 12.3
Latent correlation between the five domains

	Reading		Science		Interest		Support	
	r	SE	r	SE	r	SE	r	SE
Mathematics								
OECD	0.80	0.0009	0.89	0.0006	-0.09	0.0022	0.19	0.0025
All	0.79	0.0008	0.88	0.0004	-0.21	0.0027	0.14	0.0020
Reading								
OECD			0.84	0.0008	-0.09	0.0016	0.22	0.0014
All			0.83	0.0007	-0.18	0.0022	0.18	0.0019
Science								
OECD					-0.06	0.0022	0.25	0.0022
All					-0.19	0.0018	0.19	0.0019
Interest								
OECD							0.60	0.0014
All							0.60	0.0009

Science scales

A five-dimensional scaling was performed on the achievement data, consisting of mathematics, reading, and the three competency scales of science: *explaining phenomena scientifically*, *identifying scientific issues* and *using scientific evidence*. Responses within these domains were included in the scaling model to improve the estimation of posterior distributions of the science competency scales. The correlations between the scales are given in Table 12.4.

Table 12.4
Latent correlation between the science scales

	Identifying scientific issues		Using scientific evidence	
	r	SE	r	SE
Explaining phenomena scientifically				
OECD	0.90	0.0005	0.93	0.0002
All	0.89	0.0003	0.93	0.0002
Identifying scientific issues				
OECD			0.91	0.0002
All			0.90	0.0003
Using scientific evidence				
OECD	0.90	0.0005	0.93	0.0002
All	0.89	0.0003	0.93	0.0002



SCALING OUTCOMES

The procedures for the national and international scaling are outlined in Chapter 9 and are not reiterated here.

National item deletions

The items were first scaled by country and their fit was considered at the national level, as was the consistency of the item parameter estimates across countries. Consortium staff then adjudicated items, considering the items' functioning both within and across countries in detail. Those items considered to be dodgy (see Chapter 9) were then reviewed in consultation with NPMs. The consultations resulted in the deletion of a number of items at the national level.

At the international level, five science items were deleted from scaling (*S421Q02*, *S456Q01T*, *S456Q02*, *S426Q01* and *S508Q04*). Of these five items, *S421Q02*, *S456Q01T*, and *S456Q02* were deleted because they were misconceived by students, not because of an error in the source version. For this reason, they were added to the public student questionnaire database, but excluded from the data files with responses to cognitive items. The nationally deleted items are listed in Table 12.5. All deleted items were recoded as not applicable and were excluded from both international scaling and generating plausible values

Table 12.5
Items deleted at the national level

Item	Country
M273Q01	Lithuania (booklet 13)
M302Q01T	Turkey
M302Q02	Turkey
M420Q01T	Korea
M442Q02	Iceland (booklet 7)
M464Q01T	Croatia
M800Q01	Uruguay
R055Q01	Lithuania (booklets 6 and 11)
R102Q04A	Israel (booklets 2 and 12 of Arabic-language version), Korea
R102Q05	Chile, Israel (booklets 2 and 12 of Arabic-language version), Tunisia
R102Q07	Israel (booklets 2 and 12 of Arabic-language version), Luxembourg, Austria
R111Q02B	Lithuania, Slovak Republic
R219Q01E	Turkey
R219Q01T	Turkey
R220Q02B	Brazil (booklet 12), Denmark
R220Q06	Estonia (Russian-language version)
R227Q02	Azerbaijan (booklet 13 of Azerbaijani-language version)
S131Q04T	Mexico
S268Q02	Norway
S437Q03	Russian Federation
S447Q02	Switzerland (Italian-language version)
S447Q03	Hungary, Slovak Republic
S465Q04	Switzerland (Italian-language version)
S466Q01	Japan
S495Q03	Azerbaijan (booklet 2 of Azerbaijani-language version)
S495Q04T	Switzerland (Italian-language version), Poland
S519Q01	Azerbaijan (Russian-language version)
S519Q03	Sweden
S524Q07	Norway

Table 12.6
Final reliability of the PISA scales

Domain	Reliability
Mathematics	0.892
Reading	0.891
Science	0.920
Explaining phenomena scientifically	0.912
Identifying scientific issues	0.904
Using scientific evidence	0.923
Interest	0.892
Support	0.818



Table 12.7
National reliabilities for the main domains

	Mathematics	Reading	Science	Interest	Support
OECD					
Australia	0.88	0.88	0.92	0.90	0.83
Austria	0.90	0.93	0.94	0.89	0.82
Belgium	0.92	0.91	0.94	0.89	0.80
Canada	0.88	0.88	0.91	0.91	0.83
Czech Republic	0.90	0.91	0.92	0.89	0.77
Denmark	0.88	0.90	0.92	0.90	0.80
Finland	0.87	0.87	0.90	0.90	0.82
France	0.90	0.89	0.93	0.89	0.80
Germany	0.92	0.92	0.93	0.89	0.83
Greece	0.86	0.87	0.91	0.88	0.80
Hungary	0.91	0.90	0.91	0.88	0.80
Iceland	0.88	0.89	0.92	0.91	0.85
Ireland	0.89	0.88	0.92	0.89	0.80
Italy	0.91	0.92	0.93	0.88	0.85
Japan	0.89	0.86	0.91	0.91	0.86
Korea	0.89	0.89	0.91	0.90	0.82
Luxembourg	0.89	0.89	0.93	0.89	0.82
Mexico	0.87	0.86	0.89	0.85	0.80
Netherlands	0.93	0.92	0.93	0.90	0.78
New Zealand	0.89	0.90	0.93	0.91	0.84
Norway	0.88	0.88	0.91	0.91	0.85
Poland	0.89	0.87	0.91	0.88	0.79
Portugal	0.89	0.90	0.92	0.87	0.83
Slovak Republic	0.90	0.89	0.92	0.87	0.80
Spain	0.90	0.91	0.92	0.89	0.82
Sweden	0.88	0.89	0.92	0.90	0.83
Switzerland	0.90	0.90	0.93	0.89	0.81
Turkey	0.88	0.85	0.91	0.90	0.85
United Kingdom	0.89	0.88	0.93	0.90	0.82
United States	0.90	m	0.93	0.91	0.83
Partners					
Argentina	0.86	0.85	0.90	0.86	0.78
Azerbaijan	0.85	0.80	0.84	0.83	0.80
Brazil	0.87	0.85	0.90	0.86	0.79
Bulgaria	0.88	0.89	0.92	0.90	0.83
Chile	0.88	0.83	0.90	0.88	0.79
Colombia	0.85	0.81	0.87	0.83	0.77
Croatia	0.88	0.91	0.91	0.88	0.77
Estonia	0.88	0.91	0.91	0.87	0.78
Hong Kong-China	0.89	0.88	0.92	0.91	0.83
Indonesia	0.86	0.85	0.87	0.83	0.81
Israel	0.88	0.87	0.91	0.91	0.82
Jordan	0.86	0.86	0.90	0.85	0.83
Kyrgyzstan	0.83	0.83	0.85	0.81	0.77
Latvia	0.87	0.87	0.90	0.86	0.77
Liechtenstein	0.88	0.91	0.93	0.88	0.81
Lithuania	0.89	0.88	0.92	0.88	0.80
Macao-China	0.85	0.81	0.89	0.89	0.79
Montenegro	0.85	0.88	0.89	0.88	0.77
Qatar	0.85	0.86	0.88	0.89	0.87
Romania	0.87	0.86	0.90	0.86	0.83
Russian Federation	0.84	0.82	0.89	0.85	0.79
Serbia	0.87	0.87	0.90	0.88	0.78
Slovenia	0.90	0.93	0.93	0.90	0.80
Chinese Taipei	0.90	0.87	0.92	0.91	0.81
Thailand	0.84	0.85	0.88	0.85	0.86
Tunisia	0.86	0.83	0.87	0.85	0.83
Uruguay	0.86	0.84	0.90	0.86	0.76



Table 12.8
National reliabilities for the science scales

	Explaining phenomena scientifically	Identifying scientific issues	Using scientific evidence
OECD	Australia	0.90	0.92
	Austria	0.92	0.93
	Belgium	0.93	0.94
	Canada	0.90	0.91
	Czech Republic	0.91	0.93
	Denmark	0.91	0.92
	Finland	0.89	0.88
	France	0.92	0.93
	Germany	0.93	0.93
	Greece	0.90	0.92
	Hungary	0.91	0.92
	Iceland	0.92	0.93
	Ireland	0.92	0.93
	Italy	0.92	0.93
	Japan	0.92	0.92
	Korea	0.90	0.91
	Luxembourg	0.91	0.93
	Mexico	0.89	0.90
	Netherlands	0.92	0.93
	New Zealand	0.92	0.92
	Norway	0.90	0.92
	Poland	0.91	0.92
	Portugal	0.92	0.92
	Slovak Republic	0.91	0.93
	Spain	0.91	0.93
	Sweden	0.92	0.93
	Switzerland	0.93	0.93
Turkey	0.90	0.92	
United Kingdom	0.92	0.93	
United States	0.92	0.93	
Partners	Argentina	0.89	0.91
	Azerbaijan	0.85	0.87
	Brazil	0.89	0.92
	Bulgaria	0.91	0.92
	Chile	0.89	0.91
	Colombia	0.88	0.90
	Croatia	0.91	0.92
	Estonia	0.89	0.91
	Hong Kong-China	0.91	0.92
	Indonesia	0.84	0.85
	Israel	0.91	0.92
	Jordan	0.88	0.90
	Kyrgyzstan	0.83	0.89
	Latvia	0.89	0.92
	Liechtenstein	0.92	0.93
	Lithuania	0.91	0.92
	Macao-China	0.88	0.89
	Montenegro	0.87	0.88
	Qatar	0.87	0.87
	Romania	0.88	0.90
	Russian Federation	0.88	0.90
	Serbia	0.88	0.91
	Slovenia	0.91	0.93
	Chinese Taipei	0.91	0.91
	Thailand	0.88	0.89
	Tunisia	0.86	0.90
	Uruguay	0.90	0.92



International scaling

The international scaling was performed on the calibration data set of 15 000 students (500 randomly selected students from each of the 30 countries). The item parameter estimates from this scaling are reported in Appendix 1. The item parameters were estimated using four separate one-dimensional models. As in previous cycles, a booklet facet was used in the item response model.

Generating student scale scores

Applying the conditioning approach described in Chapter 9 and anchoring all of the item parameters at the values obtained from the international scaling, plausible values were generated for all sampled students. Table 12.6 gives the reliabilities at the international level for the generated scale scores. The increase in reliability of the results reported in Table 12.6 over those presented in Table 12.2 is due to the use of multi-dimensional scaling and conditioning.

TEST LENGTH ANALYSIS

Table 12.9 shows the number of missing responses and the number of missing responses recoded as not reached, by booklet. A response is coded as missing if the student was expected to answer a question, but no response was actually provided. All consecutive missing values clustered at the end of a test session were replaced by the non-reached code, except for the first value of the missing series, which is coded as missing (see Chapter 18).

Table 12.9
Average number of not-reached items and missing items by booklet

Booklet	Missing		Not Reached	
	Weighted	Unweighted	Weighted	Unweighted
1	4.44	4.51	3.32	2.38
2	5.40	5.46	1.84	1.50
3	4.80	4.82	1.68	1.38
4	5.33	5.41	1.71	1.38
5	5.22	5.39	5.39	4.10
6	5.48	5.94	3.81	2.95
7	5.87	5.86	1.97	1.63
8	5.65	5.70	2.62	1.97
9	5.74	6.01	2.70	2.18
10	5.46	5.59	3.48	2.88
11	5.83	6.21	3.30	2.64
12	6.07	6.20	3.37	2.66
13	5.05	5.25	2.03	1.58
UH	5.27	4.04	0.66	0.42
Total	5.41	5.56	2.86	2.25

Table 12.10 shows this information by country over all booklets. The average number of not reached items differs from one country to another. Generally, countries with higher averages of not-reached items also have higher averages of missing data. Table 12.10 provides the percentage distribution of not-reached items per booklet. The percentage of students who reached the last item ranges from 76 to 87% when using weighted data and 79 to 90% when using unweighted data (*i.e.*, the percentages of students with zero not-reached items).



Table 12.10
Average number of not-reached items and missing items by country

	Missing		Not Reached		
	Weighted	Unweighted	Weighted	Unweighted	
OECD	Australia	3.27	3.52	0.70	0.85
	Austria	4.70	4.62	0.41	0.38
	Belgium	4.14	3.89	0.99	0.90
	Canada	2.57	2.95	1.04	0.92
	Czech Republic	5.68	4.83	0.54	0.50
	Denmark	5.21	5.21	1.29	1.29
	Finland	2.50	2.50	0.59	0.58
	France	5.60	5.56	1.88	1.84
	Germany	4.83	4.81	0.67	0.66
	Greece	6.93	6.91	1.93	1.89
	Hungary	4.62	4.40	0.79	0.69
	Iceland	4.23	4.21	1.24	1.21
	Ireland	3.44	3.37	0.72	0.73
	Italy	7.47	6.88	1.74	1.54
	Japan	5.31	5.22	0.91	0.89
	Korea	2.82	2.85	0.32	0.32
	Luxembourg	5.45	5.40	0.88	0.87
	Mexico	4.14	3.89	5.42	4.78
	Netherlands	1.26	1.15	0.18	0.17
	New Zealand	3.30	3.21	0.94	0.88
	Norway	5.78	5.78	1.36	1.30
	Poland	4.89	4.70	0.84	0.82
	Portugal	5.74	5.46	1.46	1.42
	Slovak Republic	6.00	5.82	0.75	0.71
	Spain	5.96	4.96	1.70	1.35
	Sweden	4.93	4.92	1.29	1.33
Switzerland	4.16	4.24	0.77	0.76	
Turkey	6.81	6.45	1.62	1.60	
United Kingdom	4.16	4.10	1.11	0.87	
United States	2.78	2.87	0.46	0.43	
Partners	Argentina	11.23	10.87	9.97	9.34
	Azerbaijan	13.37	13.17	0.26	0.26
	Brazil	9.29	9.69	6.04	6.42
	Bulgaria	11.45	11.11	3.09	2.87
	Chile	7.67	7.44	5.09	4.98
	Colombia	5.97	5.92	12.72	12.47
	Croatia	3.91	3.93	0.75	0.75
	Estonia	3.32	3.33	0.56	0.58
	Hong Kong-China	2.39	2.23	0.74	0.71
	Indonesia	5.29	5.96	4.09	4.36
	Israel	9.73	9.62	3.16	3.16
	Jordan	6.75	6.17	3.37	3.19
	Kyrgyzstan	16.67	16.43	12.37	11.84
	Latvia	4.03	3.91	1.41	1.42
	Liechtenstein	3.71	3.73	0.42	0.42
	Lithuania	5.36	5.41	1.04	1.10
	Macao-China	3.17	3.31	1.58	1.54
	Montenegro	12.04	12.19	1.26	1.33
	Qatar	11.40	11.27	3.22	3.15
	Romania	6.95	7.29	1.09	1.24
	Russian Federation	6.18	6.09	4.33	4.26
	Serbia	10.69	10.58	2.01	1.91
	Slovenia	5.00	6.17	0.31	0.43
	Chinese Taipei	3.16	2.76	0.56	0.49
	Thailand	4.36	4.32	2.07	2.12
	Tunisia	8.56	8.51	6.10	6.02
Uruguay	10.65	10.15	7.79	7.31	



Table 12.11
Distribution of not-reached items by booklet

Number of not-reached items	Booklet														
	1	2	3	4	5	6	7	8	9	10	11	12	13	UH	
	Weighted percentage														
0	81.6	82.4	76.2	80.5	77.9	82.3	80.2	86.8	76.9	82.3	82.6	80.5	84.3	84.2	
1	1.8	0.6	7.2	2.3	1.0	0.8	0.5	0.4	1.7	1.1	0.3	0.3	0.6	3.5	
2	0.7	0.5	0.8	3.4	0.7	0.3	1.7	0.9	3.4	1.8	0.4	1.7	0.6	5.6	
3	1.1	0.5	2.4	3.0	1.3	0.3	2.5	0.4	1.5	0.4	0.2	1.8	0.6	1.2	
4	0.3	3.9	4.0	1.1	1.1	0.5	0.6	0.8	0.8	0.5	0.1	0.5	1.4	0.9	
5	1.2	0.1	0.7	1.7	0.5	0.1	0.8	0.4	3.7	0.6	1.1	1.1	0.4	1.4	
6	0.1	0.3	0.6	0.2	0.1	0.6	2.7	0.2	0.8	0.5	2.0	1.5	0.6	0.4	
7	1.0	1.8	0.3	0.5	0.1	0.5	1.3	0.5	0.7	1.2	0.6	0.3	1.6	0.1	
8	1.0	1.2	0.5	0.3	0.7	0.6	0.6	0.1	1.3	0.1	0.9	0.3	0.8	0.1	
>8	11.5	8.6	7.2	7.0	16.5	14.1	9.0	9.6	9.5	11.3	11.8	12.0	9.1	2.4	
	Unweighted percentage														
0	85.0	84.6	79.9	82.3	81.7	85.4	82.6	89.8	79.2	84.4	85.7	82.6	86.8	88.7	
1	1.8	0.6	5.6	2.6	1.0	1.0	0.7	0.3	1.7	1.1	0.3	0.4	0.6	2.8	
2	0.5	0.4	0.8	3.0	0.7	0.4	1.6	0.8	3.3	1.9	0.4	1.9	0.6	3.5	
3	1.0	0.6	2.5	3.2	1.3	0.2	2.7	0.4	1.4	0.3	0.1	2.2	0.6	1.0	
4	0.4	4.0	3.5	0.8	1.2	0.4	0.6	0.8	0.7	0.5	0.1	0.6	1.4	0.4	
5	1.1	0.2	0.8	1.6	0.4	0.1	1.0	0.2	3.9	0.7	1.0	0.7	0.5	1.9	
6	0.1	0.3	0.7	0.2	0.1	0.5	2.1	0.1	0.7	0.5	1.7	1.7	0.6	0.5	
7	0.8	1.6	0.2	0.5	0.1	0.5	1.3	0.2	0.5	0.9	0.5	0.3	1.5	0.1	
8	0.7	1.0	0.4	0.3	0.6	0.4	0.5	0.1	1.1	0.2	0.7	0.3	0.6	0.1	
>8	8.6	6.6	5.4	5.4	12.9	11.0	7.0	7.1	7.6	9.5	9.5	9.3	7.0	1.1	

Table 12.12
Estimated booklet effects on the PISA scale

Booklet	Domains				
	Mathematics	Reading	Science	Interest	Support
1			-3.1	0.6	2.6
2		10.1	-20.0	4.9	-3.3
3	2.4		-20.5	-10.5	1.1
4	11.8		-6.3	-10.7	-8.4
5	4.2		1.6	-4.5	-1.5
6		-1.2	6.7	-5.2	-1.1
7		-17.6	-19.9	13.9	3.2
8	-10.8		17.6	3.3	-3.3
9	-10.1	40.4	0.2	-0.6	4.1
10	-11.7		21.4	-1.2	-2.0
11	-13.9	19.2	12.1	-2.2	1.7
12	-5.1	-33.6	20.4	-2.5	5.9
13	19.5	-17.2	-10.4	14.6	0.9

BOOKLET EFFECTS

The booklet parameters that are described in Chapter 9 are reported in Table 12.13. The booklet effects are the amount that must be added to the proficiencies of students who responded to each booklet. That is, a positive value indicates a booklet that was harder than the average while a negative value indicates a booklet that was easier than the average. Since the booklet effects are deviations from an average they sum to zero for each domain. Table 12.13 shows the booklet effects after transformation to the PISA scales.

Table 12.13
Estimated booklet effects in logits

Booklet	Domains				
	Mathematics	Reading	Science	Interest	Support
1			-0.033	0.007	0.023
2	0.031	0.126	-0.214	0.055	-0.029
3	0.152		-0.220	-0.118	0.010
4	0.054		-0.068	-0.120	-0.073
5			0.017	-0.050	-0.013
6		-0.015	0.072	-0.058	-0.010
7	0.175	-0.220	-0.213	0.156	0.028
8	-0.139		0.189	0.037	-0.029
9	-0.130	0.506	0.002	-0.007	0.036
10	-0.150		0.229	-0.013	-0.017
11	-0.178	0.240	0.130	-0.025	0.015
12	-0.065	-0.421	0.219	-0.028	0.051
13	0.250	-0.216	-0.112	0.163	0.008



Table 12.14 [Part 1/2]
Variance in mathematics booklet means

	Expected mean	Booklet 1		Booklet 2		Booklet 3		Booklet 4		Booklet 5		Booklet 6		Booklet 7	
		Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²
OECD															
Australia	520	521	0.14	526	2.99	515	1.82	517	0.90	521	0.25	521	0.18	524	1.49
Austria	507	505	0.26	513	1.35	515	1.65	507	0.00	503	0.65	506	0.04	509	0.08
Belgium	526	526	0.00	520	2.51	531	1.17	521	2.03	527	0.03	526	0.01	529	0.60
Canada	527	528	0.02	529	0.11	528	0.02	532	1.63	536	5.17	529	0.28	529	0.11
Czech Republic	516	513	0.15	521	1.11	520	0.62	513	0.20	515	0.00	515	0.00	511	0.58
Denmark	513	508	0.92	519	1.01	521	2.01	510	0.43	511	0.17	518	0.70	514	0.02
Finland	548	550	0.28	570	19.88	550	0.25	546	0.10	546	0.07	553	1.58	550	0.41
France	496	492	0.52	491	0.63	489	1.73	499	0.51	496	0.00	502	1.93	493	0.44
Germany	509	516	1.71	504	1.14	515	0.94	512	0.44	505	0.48	505	0.60	511	0.11
Greece	460	456	0.33	457	0.18	448	5.35	464	0.74	456	0.30	469	2.50	452	1.72
Hungary	491	494	0.38	501	3.97	489	0.11	499	2.97	485	1.01	490	0.01	485	1.40
Iceland	506	510	0.33	495	3.96	501	0.74	493	5.75	505	0.03	500	1.34	510	0.48
Ireland	502	500	0.19	515	4.93	490	4.31	515	9.59	501	0.01	499	0.43	492	3.64
Italy	461	456	2.04	472	6.78	459	0.43	463	0.27	460	0.08	460	0.12	456	2.84
Japan	524	524	0.00	496	26.38	540	12.49	517	2.16	525	0.06	527	0.53	527	0.41
Korea	548	553	0.91	540	2.11	545	0.23	551	0.28	545	0.18	545	0.16	552	0.47
Luxembourg	490	486	0.41	473	13.72	497	1.87	485	0.96	488	0.07	494	0.74	500	3.82
Mexico	407	403	0.71	404	0.32	398	4.81	397	4.27	407	0.00	412	1.14	393	9.21
Netherlands	535	536	0.01	533	0.14	538	0.47	531	0.83	539	0.54	533	0.23	541	1.83
New Zealand	522	518	0.64	526	0.41	523	0.04	529	2.21	515	1.73	523	0.10	527	1.06
Norway	490	486	0.68	488	0.10	497	1.61	489	0.02	484	1.13	484	1.35	498	2.20
Poland	495	494	0.05	496	0.03	499	0.50	496	0.01	502	1.93	496	0.00	489	1.94
Portugal	466	466	0.00	468	0.10	464	0.19	475	3.80	467	0.00	458	2.37	471	0.71
Slovak Republic	494	494	0.00	489	0.77	495	0.03	500	1.42	492	0.21	495	0.12	501	1.58
Spain	480	482	0.18	482	0.32	483	0.48	477	0.56	482	0.24	477	0.22	471	4.63
Sweden	502	506	0.57	502	0.00	508	0.77	504	0.12	497	1.13	499	0.24	502	0.00
Switzerland	529	530	0.05	516	9.18	536	2.18	527	0.22	537	2.93	531	0.09	534	0.84
Turkey	424	425	0.03	439	4.92	433	1.88	424	0.00	431	1.08	424	0.00	414	2.21
United Kingdom	496	492	0.65	499	0.49	499	1.02	502	2.11	498	0.19	498	0.26	490	1.65
United States	475	481	1.16	471	0.53	466	2.54	480	0.78	476	0.04	470	0.62	470	0.71
OECD average	499	498	0.00	499	0.00	500	0.05	499	0.01	498	0.00	499	0.00	498	0.01
Partners															
Argentina	382	380	0.06	389	0.64	368	2.59	385	0.17	384	0.06	374	1.17	355	6.72
Azerbaijan	474	475	0.09	454	56.07	510	164.86	442	152.10	477	0.58	484	4.44	490	26.52
Brazil	372	366	1.39	369	0.26	362	2.57	359	4.25	376	0.40	372	0.00	365	2.22
Bulgaria	414	416	0.08	401	2.88	403	1.79	414	0.00	410	0.25	415	0.00	414	0.01
Chile	413	411	0.03	431	11.89	398	5.47	415	0.33	408	0.50	414	0.04	398	5.28
Chinese Taipei	549	548	0.03	546	0.50	561	4.36	543	1.61	550	0.01	549	0.01	546	0.40
Colombia	371	379	1.18	362	1.99	351	6.63	368	0.35	363	1.63	367	0.44	348	13.74
Croatia	468	468	0.01	466	0.10	478	5.73	462	1.69	472	1.60	466	0.06	468	0.00
Estonia	514	516	0.08	503	5.21	520	1.68	521	1.57	515	0.00	519	0.75	505	3.77
Hong Kong-China	548	551	0.35	554	2.08	548	0.01	544	0.51	548	0.00	547	0.04	550	0.28
Indonesia	391	396	0.62	393	0.12	386	0.45	390	0.00	390	0.02	388	0.12	390	0.02
Israel	443	447	0.36	438	0.36	436	0.68	435	2.05	434	1.19	443	0.00	429	3.87
Jordan	384	384	0.00	392	2.79	385	0.02	385	0.01	382	0.16	379	0.91	364	16.38
Kyrgyzstan	312	310	0.15	316	0.64	282	28.46	315	0.51	313	0.05	308	0.48	300	5.08
Latvia	486	487	0.01	484	0.14	484	0.24	483	0.24	490	0.55	493	1.18	473	5.23
Liechtenstein	525	518	0.12	523	0.04	523	0.01	539	0.40	551	1.86	497	1.68	514	0.30
Lithuania	486	492	1.20	496	3.95	494	2.55	480	1.95	484	0.13	486	0.00	473	7.66
Macao-China	525	522	0.27	527	0.06	531	1.48	533	2.48	524	0.07	527	0.05	516	2.91
Montenegro	399	405	1.16	392	2.19	404	0.69	394	0.81	406	1.26	401	0.07	402	0.49
Qatar	318	314	0.71	302	11.84	301	14.69	311	2.42	318	0.00	315	0.31	320	0.23
Romania	414	414	0.00	408	0.98	410	0.45	398	10.04	414	0.00	414	0.01	412	0.07
Russian Federation	476	473	0.21	479	0.38	469	0.97	470	1.04	473	0.18	477	0.07	465	5.51
Serbia	436	442	0.84	432	0.68	434	0.10	426	3.76	432	0.35	437	0.04	433	0.28
Slovenia	506	505	0.09	495	2.99	513	1.41	503	0.22	509	0.43	502	0.58	507	0.03
Thailand	417	416	0.03	436	16.77	412	1.58	422	1.18	418	0.06	420	0.36	408	3.74
Tunisia	368	369	0.02	392	14.45	345	13.40	362	1.10	364	0.38	363	0.56	350	7.26
Uruguay	430	436	1.14	425	0.67	401	21.24	407	14.40	422	2.60	435	0.52	398	20.78



Table 12.14 [Part 2/2]
Variance in mathematics booklet means

	Boolet 8		Boolet 9		Boolet 10		Boolet 11		Boolet 12		Boolet 13		Chi-sq (df=12)
	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	
OECD													
Australia	521	0.07	515	1.83	518	0.18	521	0.10	517	0.45	521	0.20	10.6
Austria	497	3.33	505	0.13	504	0.25	513	1.16	500	1.44	517	2.48	12.8
Belgium	522	1.15	516	6.09	523	0.66	527	0.01	540	9.07	536	6.16	29.5
Canada	527	0.03	519	5.20	519	2.98	520	3.44	523	0.86	533	2.22	22.1
Czech Republic	511	0.69	504	3.16	512	0.31	524	1.87	513	0.20	528	4.41	13.3
Denmark	515	0.25	508	0.97	513	0.00	500	4.95	519	1.29	514	0.01	12.7
Finland	540	3.01	531	10.41	546	0.05	535	10.75	544	0.93	566	12.07	59.8
France	497	0.02	504	3.26	491	1.07	496	0.01	493	0.19	498	0.16	10.5
Germany	508	0.07	508	0.09	515	1.02	514	0.71	499	2.65	507	0.34	10.3
Greece	461	0.10	466	1.99	473	8.36	474	8.21	452	2.60	442	12.22	44.6
Hungary	491	0.00	505	6.68	495	0.42	481	3.95	484	2.29	483	2.20	25.4
Iceland	512	0.94	522	11.96	503	0.27	508	0.10	511	0.55	505	0.07	26.5
Ireland	503	0.12	505	0.52	490	5.31	484	12.32	511	3.39	513	4.57	49.3
Italy	458	0.78	463	0.11	470	7.74	479	16.99	454	4.72	452	5.93	48.8
Japan	526	0.23	526	0.14	514	4.38	525	0.04	535	6.08	520	0.52	53.4
Korea	546	0.07	550	0.15	534	5.38	549	0.05	557	3.08	551	0.44	13.5
Luxembourg	495	1.28	496	1.23	489	0.02	493	0.47	488	0.14	487	0.30	25.0
Mexico	409	0.27	401	2.12	419	9.34	428	35.67	404	0.44	401	3.23	71.5
Netherlands	539	0.39	536	0.02	521	8.72	523	5.82	546	4.29	540	1.00	24.3
New Zealand	523	0.05	521	0.00	518	0.76	513	3.34	526	0.81	526	0.50	11.6
Norway	494	0.58	493	0.24	493	0.21	490	0.01	483	1.28	489	0.04	9.5
Poland	500	0.73	486	4.48	487	2.25	493	0.17	508	6.81	495	0.01	18.9
Portugal	461	1.14	466	0.01	472	0.93	473	1.13	457	2.82	464	0.11	13.3
Slovak Republic	491	0.26	509	8.03	498	0.50	496	0.23	480	8.32	482	5.51	27.0
Spain	477	0.58	478	0.31	481	0.04	484	1.38	485	1.82	480	0.00	10.8
Sweden	500	0.31	504	0.17	497	0.94	507	0.77	496	1.01	506	0.44	6.5
Switzerland	530	0.04	524	1.68	529	0.02	528	0.07	530	0.05	532	0.25	17.6
Turkey	413	2.66	413	3.04	435	3.78	419	0.68	414	2.04	427	0.23	22.6
United Kingdom	492	0.49	491	1.28	497	0.08	482	9.52	499	0.78	503	2.72	21.2
United States	484	2.12	488	6.40	465	2.31	447	23.67	487	5.08	482	1.48	47.4
OECD average	498	0.01	498	0.01	497	0.05	498	0.04	499	0.00	500	0.07	0.3
Partners													
Argentina	394	1.89	386	0.18	401	5.81	395	3.16	397	2.84	348	17.89	43.2
Azerbaijan	465	8.45	440	132.24	470	2.52	480	2.90	503	93.00	498	42.54	686.3
Brazil	365	2.78	373	0.01	390	12.84	402	57.32	366	1.60	341	30.03	115.7
Bulgaria	413	0.05	422	1.11	429	3.65	434	8.63	398	4.17	405	1.23	23.8
Chile	421	2.05	412	0.00	421	2.73	428	6.09	396	7.13	395	7.70	49.2
Chinese Taipei	549	0.02	539	4.44	545	0.60	537	4.68	561	6.17	570	11.84	34.7
Colombia	380	1.82	370	0.03	399	28.38	416	59.88	376	0.53	330	50.73	167.3
Croatia	458	4.25	454	10.60	472	1.32	481	8.57	458	4.60	471	0.86	39.4
Estonia	520	1.01	519	0.72	512	0.30	508	1.48	520	1.34	511	0.41	18.3
Hong Kong-China	548	0.00	532	12.01	539	2.82	540	2.25	555	2.29	565	11.77	34.4
Indonesia	403	2.81	401	2.29	388	0.21	379	4.08	393	0.13	386	0.59	11.5
Israel	451	1.59	454	2.25	453	3.74	448	1.08	450	1.04	428	4.20	22.4
Jordan	387	0.19	376	2.76	385	0.01	394	4.01	407	20.87	373	4.52	52.6
Kyrgyzstan	319	2.55	323	5.29	329	10.00	340	38.30	305	1.26	276	47.53	140.3
Latvia	489	0.36	490	0.39	485	0.02	479	1.87	495	3.32	487	0.01	13.6
Liechtenstein	534	0.23	510	0.72	528	0.02	535	0.29	529	0.03	525	0.00	5.7
Lithuania	477	2.44	478	2.33	483	0.40	485	0.01	495	2.71	504	9.41	34.7
Macao-China	529	0.35	525	0.00	520	1.35	509	8.17	532	1.42	531	0.94	19.6
Montenegro	394	0.78	384	10.86	399	0.00	420	20.42	392	2.18	397	0.17	41.1
Qatar	325	2.13	350	40.60	342	31.24	338	18.66	306	8.80	290	23.76	155.4
Romania	415	0.05	419	0.84	418	0.58	444	19.60	406	1.06	420	0.77	34.5
Russian Federation	485	3.75	478	0.26	484	1.60	493	9.40	478	0.18	460	6.60	30.2
Serbia	433	0.30	433	0.22	439	0.20	456	17.95	432	0.50	432	0.49	25.7
Slovenia	506	0.00	509	0.24	502	0.57	511	0.79	504	0.20	510	0.72	8.3
Thailand	405	7.68	407	5.05	436	20.96	413	1.05	410	2.16	418	0.04	60.7
Tunisia	370	0.25	353	4.29	395	24.59	386	15.84	377	2.65	328	38.62	123.4
Uruguay	445	8.02	438	2.34	450	14.42	451	20.21	450	11.00	396	25.08	142.4



Table 12.15 [Part 1/2]
Variance in reading booklet means

	Expected mean	Booklet 1		Booklet 2		Booklet 3		Booklet 4		Booklet 5		Booklet 6		Booklet 7	
		Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²
OECD															
Australia	513	514	0.16	518	2.37	509	1.10	508	1.64	514	0.14	510	0.52	516	0.71
Austria	491	493	0.08	500	2.19	490	0.04	484	1.45	495	0.35	493	0.09	483	2.54
Belgium	507	509	0.11	507	0.02	502	0.82	507	0.00	508	0.09	506	0.01	501	1.37
Canada	527	529	0.11	527	0.00	527	0.00	527	0.02	536	3.92	530	0.38	522	1.73
Czech Republic	490	485	0.55	491	0.06	495	0.99	483	1.06	489	0.01	486	0.30	489	0.00
Denmark	495	490	0.71	488	1.13	497	0.25	491	0.38	495	0.00	502	1.80	489	1.02
Finland	546	551	1.09	563	8.27	549	0.22	542	0.81	544	0.18	550	0.43	556	4.97
France	489	486	0.13	475	4.09	487	0.04	483	0.76	485	0.28	498	4.41	495	1.14
Germany	505	509	0.36	507	0.10	502	0.23	508	0.25	497	1.43	503	0.18	496	2.09
Greece	460	453	1.31	460	0.01	455	0.74	461	0.00	460	0.00	473	4.53	471	3.31
Hungary	483	483	0.03	495	4.44	481	0.06	484	0.06	477	0.98	483	0.00	479	0.32
Iceland	485	489	0.32	475	2.91	480	0.61	473	3.69	485	0.00	485	0.00	486	0.04
Ireland	517	515	0.12	528	2.93	514	0.29	518	0.03	514	0.41	508	2.33	522	0.82
Italy	469	470	0.01	457	9.42	476	1.97	473	0.85	465	0.67	470	0.02	462	2.73
Japan	498	500	0.12	484	6.25	498	0.00	501	0.23	499	0.03	494	0.76	491	1.45
Korea	556	562	1.09	555	0.01	558	0.11	551	0.91	554	0.17	554	0.21	543	5.21
Luxembourg	479	474	1.17	487	2.06	483	0.32	479	0.00	478	0.06	480	0.01	490	3.04
Mexico	412	411	0.13	378	45.75	409	0.50	408	0.62	408	0.46	429	14.21	417	1.04
Netherlands	513	512	0.03	525	5.10	512	0.00	508	0.54	518	1.16	508	0.97	516	0.46
New Zealand	521	514	1.06	529	1.78	528	1.46	524	0.27	511	3.48	518	0.33	526	0.77
Norway	485	474	2.11	478	1.14	494	2.38	490	0.47	476	1.47	473	3.63	485	0.00
Poland	508	508	0.01	509	0.07	508	0.00	515	1.28	516	1.95	503	0.83	519	4.89
Portugal	473	469	0.33	476	0.33	477	0.50	481	2.16	471	0.06	471	0.11	496	15.41
Slovak Republic	469	467	0.14	463	0.89	468	0.03	474	1.28	462	1.15	481	7.15	462	1.40
Spain	461	464	0.22	450	5.00	462	0.13	462	0.14	462	0.11	460	0.02	462	0.07
Sweden	508	509	0.01	501	0.59	500	0.83	506	0.18	512	0.44	500	1.65	513	0.87
Switzerland	499	503	0.66	491	4.48	499	0.02	502	0.29	507	3.49	496	0.59	490	5.71
Turkey	447	442	0.55	435	3.38	454	0.86	455	1.44	452	0.50	450	0.16	448	0.00
United Kingdom	496	491	0.80	492	0.66	503	2.75	498	0.16	499	0.28	496	0.01	492	0.57
United States															
OECD average	476	476	0.01	475	0.08	477	0.02	476	0.00	476	0.00	477	0.01	477	0.02
Partners															
Argentina	378	377	0.00	334	12.97	382	0.20	379	0.04	378	0.00	386	0.81	394	1.71
Azerbaijan	353	352	0.06	350	0.54	352	0.07	351	0.34	351	0.34	362	3.46	336	11.80
Brazil	393	390	0.15	373	6.33	395	0.06	383	1.61	401	1.67	398	0.95	407	6.58
Bulgaria	402	406	0.23	393	0.95	400	0.05	411	0.83	395	0.42	412	1.30	410	0.89
Chile	442	440	0.08	418	11.07	443	0.00	436	0.67	441	0.05	454	3.12	464	8.39
Chinese Taipei	497	496	0.03	501	1.14	495	0.17	497	0.02	498	0.16	490	1.91	489	3.16
Colombia	389	395	0.61	325	58.41	384	0.20	389	0.00	386	0.16	398	2.76	424	22.12
Croatia	477	480	0.45	482	0.65	478	0.02	476	0.09	478	0.04	479	0.14	482	1.21
Estonia	501	504	0.25	504	0.40	493	3.01	501	0.02	497	0.48	514	7.35	484	8.54
Hong Kong-China	535	537	0.23	544	4.72	538	0.34	533	0.19	537	0.08	520	18.12	525	4.76
Indonesia	392	396	0.37	385	1.43	397	0.67	393	0.01	393	0.00	398	0.47	399	0.83
Israel	439	446	0.74	414	10.67	433	0.47	440	0.02	430	1.15	437	0.05	433	0.56
Jordan	401	399	0.12	393	1.93	400	0.00	403	0.21	401	0.02	409	1.95	399	0.05
Kyrgyzstan	284	285	0.01	267	7.25	282	0.13	286	0.12	289	0.55	291	1.22	309	18.86
Latvia	479	483	0.33	476	0.16	479	0.00	477	0.10	484	0.58	490	3.11	488	1.74
Liechtenstein	511	503	0.23	504	0.16	508	0.02	512	0.00	535	1.69	476	2.87	497	0.54
Lithuania	470	477	1.62	490	10.74	463	1.39	473	0.18	464	1.04	472	0.07	472	0.07
Macao-China	492	491	0.10	495	0.28	495	0.27	490	0.15	492	0.02	495	0.35	488	1.17
Montenegro	392	397	0.79	390	0.21	391	0.01	390	0.08	397	0.86	394	0.19	380	4.79
Qatar	313	305	1.44	308	0.65	306	1.84	313	0.00	317	0.53	307	0.88	309	0.41
Romania	397	400	0.18	393	0.37	400	0.28	392	0.41	392	0.30	397	0.00	407	3.45
Russian Federation	440	446	1.04	417	12.27	442	0.03	438	0.14	433	2.11	452	2.50	454	7.60
Serbia	401	403	0.07	405	0.49	402	0.03	398	0.26	401	0.00	411	3.68	400	0.03
Slovenia	498	495	0.22	487	3.40	499	0.12	492	0.90	501	0.47	504	2.21	477	15.31
Thailand	417	417	0.00	411	1.24	411	1.58	416	0.07	421	0.67	423	2.36	426	4.41
Tunisia	382	382	0.01	365	6.04	386	0.24	375	1.13	379	0.25	395	5.38	386	0.54
Uruguay	414	416	0.07	360	50.73	410	0.24	409	0.67	411	0.20	429	4.35	427	2.43



Table 12.15 [Part 2/2]
Variance in reading booklet means

	Boolet 8		Boolet 9		Boolet 10		Boolet 11		Boolet 12		Boolet 13		Chi-sq (df=12)
	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	
OECD													
Australia	518	1.40	518	1.99	511	0.36	512	0.01	510	0.81	509	0.77	12.0
Austria	492	0.01	502	2.36	493	0.03	498	1.17	479	4.72	490	0.02	15.1
Belgium	504	0.38	506	0.01	503	0.71	513	1.47	511	1.01	510	0.46	6.5
Canada	529	0.13	526	0.16	523	0.66	528	0.01	520	2.84	530	0.64	10.6
Czech Republic	486	0.22	487	0.19	490	0.00	494	0.29	488	0.06	498	2.36	6.1
Denmark	495	0.00	509	5.05	494	0.00	502	1.69	479	6.96	498	0.47	19.5
Finland	543	0.51	545	0.04	549	0.43	547	0.04	536	4.94	533	9.96	31.9
France	491	0.17	482	1.21	488	0.01	489	0.01	491	0.10	490	0.03	12.4
Germany	509	0.46	517	4.54	506	0.01	498	1.23	502	0.24	507	0.11	11.2
Greece	457	0.21	435	15.62	467	1.01	445	5.82	476	7.94	464	0.30	40.8
Hungary	481	0.04	484	0.05	481	0.06	481	0.09	489	2.21	471	3.91	12.3
Iceland	485	0.00	513	24.00	483	0.10	486	0.01	475	2.62	485	0.01	34.3
Ireland	524	1.15	528	3.62	511	1.01	512	0.90	515	0.18	515	0.16	14.0
Italy	469	0.02	461	3.42	468	0.07	463	1.62	475	2.76	483	14.56	38.1
Japan	500	0.10	508	3.82	502	0.48	498	0.00	497	0.04	503	0.65	14.0
Korea	554	0.09	571	6.96	551	0.67	568	5.35	552	0.79	556	0.01	21.6
Luxembourg	480	0.00	471	1.69	482	0.20	474	0.96	477	0.13	477	0.12	9.8
Mexico	409	0.38	378	39.74	409	0.27	423	3.56	426	7.45	433	25.39	139.5
Netherlands	510	0.12	525	5.64	511	0.05	516	0.28	509	0.76	499	8.83	23.9
New Zealand	517	0.44	532	3.00	518	0.27	519	0.12	524	0.34	514	1.35	14.7
Norway	489	0.50	489	0.46	486	0.04	488	0.18	481	0.35	493	2.21	14.9
Poland	505	0.18	498	3.24	503	0.43	500	1.93	512	0.68	503	0.74	16.2
Portugal	473	0.01	449	14.20	472	0.01	456	7.89	491	10.24	461	5.38	56.6
Slovak Republic	467	0.10	478	2.36	467	0.08	459	2.36	461	1.87	474	0.79	19.6
Spain	459	0.26	461	0.01	460	0.11	463	0.41	460	0.05	465	0.83	7.4
Sweden	508	0.00	517	2.36	505	0.35	509	0.03	511	0.25	504	0.66	8.2
Switzerland	501	0.12	513	10.95	503	0.41	508	4.82	479	20.64	500	0.02	52.2
Turkey	447	0.01	436	3.12	447	0.00	454	1.33	449	0.12	444	0.30	11.8
United Kingdom	490	0.87	485	3.69	499	0.45	488	2.42	503	2.95	499	0.38	16.0
United States													
OECD average	476	0.00	477	0.03	476	0.00	476	0.00	476	0.01	477	0.01	0.2
Partners													
Argentina	367	0.60	319	25.30	370	0.51	355	6.72	413	20.38	403	8.39	77.6
Azerbaijan	353	0.01	367	11.04	352	0.04	358	1.19	337	11.51	368	7.35	47.7
Brazil	396	0.24	362	30.46	395	0.22	389	0.77	418	12.77	400	1.05	62.8
Bulgaria	404	0.04	393	1.27	397	0.24	397	0.50	400	0.08	409	0.58	7.4
Chile	442	0.00	414	13.85	437	0.77	434	1.57	461	7.33	461	7.60	54.5
Chinese Taipei	499	0.16	508	7.79	491	1.09	502	1.65	492	1.11	493	0.66	19.0
Colombia	391	0.06	309	100.96	377	1.43	388	0.00	422	25.20	428	24.28	236.2
Croatia	473	0.57	471	1.73	475	0.19	471	2.29	486	3.87	473	0.92	12.2
Estonia	507	1.30	512	3.77	500	0.00	504	0.22	489	4.91	501	0.01	30.3
Hong Kong-China	537	0.14	554	15.88	531	1.20	552	13.65	519	12.94	541	1.84	74.1
Indonesia	395	0.13	385	1.06	395	0.18	395	0.24	381	4.86	397	0.39	10.7
Israel	440	0.02	437	0.04	442	0.10	439	0.00	448	1.85	461	8.39	24.1
Jordan	402	0.08	388	5.68	398	0.24	407	1.51	406	1.23	402	0.11	13.1
Kyrgyzstan	283	0.12	245	54.62	283	0.02	276	1.88	321	37.45	285	0.01	122.2
Latvia	478	0.02	474	0.97	480	0.00	475	0.85	475	0.58	475	0.58	9.0
Liechtenstein	512	0.00	521	0.31	516	0.06	543	2.46	489	1.17	521	0.26	9.8
Lithuania	465	0.66	466	0.67	468	0.21	461	2.18	467	0.44	474	0.52	19.8
Macao-China	497	0.73	492	0.01	493	0.04	494	0.11	481	5.71	498	1.97	10.9
Montenegro	393	0.05	396	0.65	383	1.96	402	3.15	378	7.58	405	5.34	25.7
Qatar	307	0.77	295	8.77	321	2.24	312	0.00	326	9.79	331	8.52	35.9
Romania	387	1.51	380	7.00	396	0.02	379	4.43	415	7.08	411	2.89	27.9
Russian Federation	439	0.09	418	13.51	441	0.03	443	0.20	450	2.31	446	1.05	42.9
Serbia	400	0.06	396	0.90	396	0.74	405	0.59	394	1.88	401	0.00	8.7
Slovenia	498	0.03	517	15.35	496	0.10	509	5.57	478	15.51	504	2.48	61.7
Thailand	419	0.24	415	0.25	418	0.11	411	1.12	407	5.59	421	1.00	18.6
Tunisia	382	0.01	347	23.32	376	0.80	376	0.83	394	6.08	398	7.50	52.1
Uruguay	416	0.04	367	38.05	417	0.09	402	2.63	445	23.34	455	35.22	158.1



Table 12.16 [Part 1/2]
Variance in science booklet means

	Expected mean	Booklet 1		Boolet 2		Booklet 3		Boolet 4		Boolet 5		Boolet 6		Boolet 7	
		Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²
OECD															
Australia	526	531	1.32	525	0.07	517	6.85	525	0.14	530	1.25	529	0.30	516	7.59
Austria	512	512	0.00	510	0.12	510	0.12	501	3.19	507	0.79	520	1.76	501	3.66
Belgium	514	513	0.08	514	0.00	511	1.02	513	0.21	520	2.27	514	0.01	523	4.92
Canada	534	535	0.04	536	0.12	525	6.72	532	0.27	545	7.60	534	0.01	532	0.19
Czech Republic	518	520	0.18	516	0.18	519	0.08	511	2.25	517	0.04	516	0.17	532	5.96
Denmark	496	488	1.60	496	0.00	502	1.22	499	0.36	496	0.00	516	12.61	488	1.64
Finland	564	565	0.10	553	3.86	563	0.04	563	0.01	570	1.28	581	11.15	571	2.55
France	495	484	3.83	513	8.90	500	1.17	493	0.14	498	0.33	491	0.82	510	8.64
Germany	522	530	2.77	522	0.02	519	0.32	522	0.03	516	0.94	523	0.05	505	7.32
Greece	473	473	0.03	479	0.91	493	12.87	479	1.34	471	0.19	469	0.63	496	14.27
Hungary	504	506	0.20	507	0.39	500	0.62	501	0.33	501	0.38	510	1.43	501	0.41
Iceland	491	490	0.05	479	4.41	483	1.99	487	0.51	487	0.50	492	0.04	495	0.50
Ireland	508	506	0.19	505	0.17	505	0.34	510	0.16	496	7.69	508	0.00	474	32.09
Italy	476	472	1.19	490	15.61	486	7.28	477	0.18	471	2.15	479	0.92	476	0.00
Japan	531	531	0.00	528	0.21	522	3.50	517	9.82	531	0.00	523	2.13	533	0.22
Korea	522	526	0.74	498	20.85	500	23.17	519	0.47	521	0.01	523	0.07	523	0.06
Luxembourg	486	479	2.11	489	0.21	490	0.58	485	0.06	484	0.20	494	1.78	488	0.13
Mexico	410	399	7.15	419	4.60	416	1.87	409	0.12	410	0.01	402	3.43	417	2.29
Netherlands	529	534	1.34	527	0.14	525	1.04	525	0.54	539	3.96	532	0.52	529	0.00
New Zealand	531	518	4.36	530	0.00	527	0.36	534	0.43	520	3.93	536	0.99	527	0.28
Norway	487	476	3.82	482	0.66	499	5.07	498	3.47	481	0.92	487	0.01	498	3.51
Poland	498	495	0.24	491	1.77	498	0.00	506	2.42	509	5.12	491	1.82	503	1.27
Portugal	475	481	1.15	481	1.42	472	0.22	476	0.14	475	0.00	459	7.97	473	0.07
Slovak Republic	490	492	0.35	480	3.18	492	0.19	487	0.21	489	0.01	486	0.86	498	1.76
Spain	488	487	0.04	490	0.07	490	0.06	482	3.03	492	0.72	478	3.87	499	6.11
Sweden	503	510	1.22	498	0.69	505	0.03	503	0.01	501	0.23	505	0.08	518	7.60
Switzerland	511	512	0.10	506	1.92	505	1.54	516	1.10	519	2.87	514	0.26	506	1.53
Turkey	425	427	0.16	424	0.03	430	1.07	434	4.24	431	1.89	413	3.45	416	2.36
United Kingdom	516	514	0.20	515	0.04	530	15.08	516	0.00	513	0.40	517	0.04	504	5.37
United States	490	492	0.17	489	0.04	497	1.10	496	1.16	486	0.36	477	3.40	478	3.60
OECD average	501	500	0.02	500	0.04	501	0.00	501	0.00	501	0.00	501	0.00	501	0.00
Partners															
Argentina	392	392	0.00	418	10.15	410	5.71	406	4.52	384	0.75	363	16.00	406	1.96
Azerbaijan	381	395	10.46	380	0.11	393	9.50	383	0.25	367	16.15	405	28.49	383	0.21
Brazil	391	379	4.57	410	15.18	421	45.49	402	5.42	396	1.27	377	8.55	388	0.18
Bulgaria	434	433	0.05	435	0.00	441	0.88	440	0.56	428	0.56	429	0.37	450	3.93
Chile	439	442	0.37	455	9.13	448	2.49	446	2.49	429	2.56	421	8.87	438	0.00
Chinese Taipei	532	533	0.03	531	0.07	526	1.83	517	10.04	528	0.72	519	8.04	543	4.84
Colombia	389	388	0.04	409	10.42	416	13.02	411	18.29	394	0.71	363	26.20	405	5.07
Croatia	493	496	0.42	498	0.78	495	0.18	485	3.19	484	5.36	490	0.43	483	6.70
Estonia	531	533	0.09	531	0.01	537	1.33	527	0.66	525	1.50	550	10.29	517	6.96
Hong Kong-China	541	546	1.35	540	0.06	533	3.96	535	1.88	539	0.24	538	0.66	544	0.24
Indonesia	394	395	0.02	398	0.44	414	11.82	406	2.37	392	0.09	376	5.83	400	0.51
Israel	453	450	0.13	461	1.94	470	4.96	451	0.15	450	0.18	458	0.68	475	9.06
Jordan	423	422	0.03	434	4.96	427	1.14	422	0.01	420	0.21	403	13.35	415	2.48
Kyrgyzstan	324	324	0.00	347	21.36	343	16.92	344	24.56	318	1.28	295	35.63	334	6.35
Latvia	490	486	0.27	482	1.26	490	0.03	492	0.15	487	0.21	493	0.23	493	0.27
Liechtenstein	524	511	0.36	530	0.15	511	0.30	538	0.48	544	1.03	488	2.74	512	0.28
Lithuania	488	494	1.49	487	0.06	489	0.04	492	0.65	481	1.75	492	0.64	495	1.81
Macao-China	512	508	0.65	500	4.39	513	0.08	512	0.05	507	1.11	505	1.68	519	3.88
Montenegro	412	421	2.34	417	1.13	415	0.34	408	0.98	408	0.61	401	5.41	413	0.00
Qatar	350	336	9.77	370	20.51	354	1.37	357	3.65	352	0.44	333	12.56	350	0.00
Romania	418	422	0.41	434	8.18	423	0.99	409	4.05	411	0.91	409	2.86	413	0.97
Russian Federation	479	475	0.67	498	11.49	494	4.99	486	1.50	475	1.22	466	3.39	487	2.60
Serbia	435	447	3.77	435	0.01	436	0.00	431	0.80	426	2.80	431	0.92	438	0.37
Slovenia	522	523	0.04	504	8.39	523	0.00	503	7.83	523	0.01	524	0.14	529	1.22
Thailand	421	412	8.28	421	0.04	438	19.19	432	7.12	428	2.78	413	3.34	430	4.33
Tunisia	385	396	4.28	409	18.40	418	44.11	399	8.60	377	3.23	366	13.71	388	0.18
Uruguay	429	429	0.00	459	36.38	448	14.95	425	0.40	422	1.58	413	6.14	445	6.28



Table 12.16 [Part 2/2]
Variance in science booklet means

	Boolet 8		Boolet 9		Boolet 10		Boolet 11		Boolet 12		Boolet 13		Chi-sq (df=12)
	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	
OECD													
Australia	536	3.68	529	0.35	529	0.30	534	2.59	529	0.45	519	2.53	27.4
Austria	517	0.58	517	0.76	517	0.73	513	0.02	511	0.02	522	2.72	14.4
Belgium	508	2.32	505	6.33	507	3.20	518	0.84	517	0.33	526	8.07	29.6
Canada	534	0.00	530	1.26	532	0.25	539	1.04	535	0.03	538	0.94	18.5
Czech Republic	520	0.08	511	1.41	517	0.05	516	0.08	511	1.72	531	5.25	17.4
Denmark	509	4.58	490	0.81	498	0.16	484	4.55	483	4.93	496	0.00	32.5
Finland	558	1.08	540	23.41	561	0.22	573	4.15	562	0.11	564	0.00	48.0
France	492	0.35	494	0.02	475	13.77	500	0.55	485	2.54	501	1.17	42.2
Germany	529	1.63	521	0.02	533	3.68	517	0.77	520	0.05	520	0.08	17.6
Greece	445	16.91	482	2.52	469	0.79	460	6.65	460	8.62	478	0.60	66.3
Hungary	504	0.00	505	0.04	513	3.26	499	1.07	509	1.38	495	3.08	12.6
Iceland	491	0.00	505	7.29	490	0.02	493	0.15	491	0.00	498	1.19	16.7
Ireland	518	2.78	530	15.83	522	5.60	513	1.01	528	10.90	494	4.80	81.5
Italy	471	1.83	480	1.70	468	3.85	464	9.18	471	1.45	476	0.02	45.3
Japan	544	5.46	544	7.65	549	11.66	532	0.03	530	0.01	524	1.60	42.3
Korea	532	3.09	511	3.70	529	1.59	538	9.85	544	19.62	525	0.46	83.7
Luxembourg	495	3.08	478	2.04	489	0.13	485	0.03	477	3.00	489	0.13	13.5
Mexico	394	16.51	422	7.82	392	18.74	411	0.08	405	1.38	429	29.20	93.2
Netherlands	536	1.37	525	0.75	533	0.58	533	0.65	521	2.77	522	1.92	15.6
New Zealand	531	0.01	538	1.30	538	1.49	528	0.15	546	7.67	523	1.48	22.5
Norway	491	0.39	488	0.02	480	1.06	482	0.63	478	2.44	484	0.29	22.3
Poland	494	0.53	495	0.24	499	0.08	493	1.09	497	0.00	499	0.14	14.7
Portugal	468	2.00	480	1.07	470	0.69	473	0.15	487	6.94	472	0.29	22.1
Slovak Republic	485	0.65	490	0.00	487	0.22	487	0.37	490	0.04	505	8.69	16.5
Spain	490	0.14	482	2.29	485	0.89	490	0.16	488	0.04	499	4.11	21.5
Sweden	504	0.01	509	1.39	503	0.01	490	7.17	497	1.39	502	0.10	19.9
Switzerland	524	7.10	500	6.23	515	0.69	513	0.24	506	1.51	515	0.42	25.5
Turkey	430	0.99	416	2.43	421	0.48	434	3.83	420	1.06	414	3.01	25.0
United Kingdom	510	1.10	519	0.34	524	2.54	508	2.78	516	0.00	506	3.00	30.9
United States	475	4.37	505	7.76	489	0.03	496	0.78	495	0.75	481	1.90	25.4
OECD average	501	0.00	501	0.02	501	0.00	501	0.00	500	0.01	502	0.03	0.1
Partners													
Argentina	369	4.48	404	1.87	362	10.26	381	2.29	378	3.12	413	7.15	68.3
Azerbaijan	389	3.62	374	4.03	360	28.00	389	3.77	373	7.34	379	0.27	112.2
Brazil	372	16.86	403	6.11	369	25.26	398	2.96	387	0.66	373	11.14	143.7
Bulgaria	426	1.28	446	2.70	421	2.61	434	0.00	421	3.41	441	0.64	17.0
Chile	423	6.33	450	3.65	419	11.69	441	0.16	427	3.99	458	9.17	60.9
Chinese Taipei	541	2.76	544	7.31	535	0.23	530	0.16	529	0.76	547	8.59	45.4
Colombia	355	28.94	396	1.44	343	38.82	385	0.69	364	17.45	417	26.10	187.2
Croatia	499	1.40	501	4.17	498	1.62	493	0.00	494	0.03	496	0.44	24.7
Estonia	539	1.88	534	0.22	535	0.40	516	7.10	537	1.05	528	0.48	32.0
Hong Kong-China	555	4.55	538	0.52	538	0.38	547	1.76	536	1.19	560	13.20	30.0
Indonesia	375	6.03	399	0.65	377	8.47	402	1.71	392	0.18	390	0.58	38.7
Israel	456	0.28	460	0.95	441	2.51	436	6.56	431	11.39	460	0.70	39.5
Jordan	408	6.99	443	17.89	415	2.06	420	0.38	432	3.74	424	0.04	53.3
Kyrgyzstan	284	78.75	336	11.20	288	46.81	337	7.18	316	2.25	323	0.00	252.3
Latvia	500	3.20	488	0.13	485	0.59	484	1.14	492	0.25	492	0.12	7.9
Liechtenstein	536	0.50	510	0.64	528	0.06	540	0.66	510	0.45	532	0.26	7.9
Lithuania	478	3.10	486	0.14	485	0.26	471	10.80	486	0.11	508	12.49	33.3
Macao-China	517	1.43	508	0.79	513	0.07	507	0.95	505	1.47	528	11.89	28.4
Montenegro	409	0.36	417	1.29	402	3.42	420	2.70	412	0.03	412	0.00	18.6
Qatar	330	22.28	368	16.29	341	5.12	366	15.79	349	0.03	334	9.71	117.5
Romania	418	0.00	433	6.81	408	2.37	424	0.47	413	0.53	421	0.18	28.7
Russian Federation	469	3.94	482	0.30	475	0.34	473	0.90	470	3.59	481	0.04	35.0
Serbia	435	0.00	442	2.42	437	0.05	430	1.33	431	0.51	443	2.14	15.1
Slovenia	510	4.12	522	0.00	534	3.99	510	4.37	525	0.23	541	15.84	46.2
Thailand	398	30.91	431	6.94	411	7.02	421	0.00	422	0.03	417	1.37	91.3
Tunisia	372	8.83	393	2.84	359	29.92	385	0.00	374	4.80	376	2.89	141.8
Uruguay	421	1.73	430	0.12	398	26.81	422	1.34	400	22.59	449	8.77	127.1



Table 12.17 [Part 1/2]
Variance in interest booklet means

	Expected mean	Booklet 1		Booklet 2		Booklet 3		Booklet 4		Booklet 5		Booklet 6		Booklet 7	
		Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²
OECD															
Australia	466	473	5.09	471	1.79	463	0.79	463	0.62	462	1.25	461	1.83	454	8.86
Austria	506	511	1.50	498	2.60	502	0.48	506	0.00	506	0.01	516	2.74	509	0.37
Belgium	501	502	0.07	498	0.51	499	0.21	507	1.89	499	0.24	505	0.77	511	6.98
Canada	470	477	3.84	477	3.39	465	1.37	464	1.58	466	0.71	467	0.29	459	5.38
Czech Republic	487	476	5.40	495	2.34	500	6.48	491	0.43	485	0.19	476	3.02	488	0.03
Denmark	463	463	0.00	454	3.02	457	1.22	464	0.06	460	0.28	482	12.25	452	4.77
Finland	449	459	3.77	452	0.46	450	0.07	448	0.08	449	0.00	433	6.30	452	0.55
France	520	520	0.02	519	0.02	522	0.17	526	1.22	514	1.27	537	11.76	514	1.34
Germany	512	513	0.04	522	4.00	512	0.00	515	0.42	512	0.00	515	0.25	503	3.16
Greece	548	543	1.26	545	0.34	535	6.99	553	1.00	556	2.64	554	1.17	559	2.44
Hungary	522	525	0.47	529	2.26	515	2.24	517	1.46	527	1.18	523	0.10	525	0.23
Iceland	467	465	0.11	472	0.53	477	2.56	458	1.65	465	0.08	446	8.87	481	3.84
Ireland	481	482	0.05	480	0.04	480	0.04	476	1.04	492	3.46	481	0.00	482	0.00
Italy	529	523	3.85	520	9.65	522	3.63	534	2.59	528	0.09	534	2.88	532	0.96
Japan	512	517	1.03	490	17.46	513	0.06	528	8.65	513	0.09	509	0.23	497	5.10
Korea	486	482	0.49	481	0.65	494	3.15	476	2.36	484	0.08	482	0.58	478	2.70
Luxembourg	515	520	1.22	519	0.66	503	6.28	511	0.61	512	0.28	518	0.33	522	2.20
Mexico	611	596	11.28	599	7.08	612	0.25	611	0.02	618	2.62	620	5.62	606	1.25
Netherlands	449	465	11.97	439	3.83	431	13.48	459	3.61	435	5.30	456	1.21	451	0.05
New Zealand	461	461	0.01	474	6.18	459	0.38	466	0.77	471	2.45	456	0.95	452	2.53
Norway	473	467	1.10	485	3.24	496	25.27	477	0.46	470	0.33	470	0.34	470	0.42
Poland	500	495	2.15	494	2.11	497	0.49	488	7.33	503	0.29	494	1.62	508	3.42
Portugal	571	565	1.62	563	2.59	564	1.97	576	1.10	562	3.09	576	1.44	576	1.45
Slovak Republic	522	522	0.00	522	0.01	517	1.07	523	0.08	521	0.05	526	0.60	528	1.62
Spain	534	526	4.34	540	2.56	540	1.65	527	2.65	532	0.22	534	0.00	531	0.69
Sweden	455	466	4.72	455	0.00	446	1.00	447	2.36	449	1.15	453	0.20	451	0.33
Switzerland	504	506	0.33	508	1.39	492	8.87	506	0.24	506	0.14	508	0.71	508	1.48
Turkey	542	535	0.99	526	4.53	565	20.84	536	0.63	536	0.67	553	4.41	559	8.05
United Kingdom	464	471	3.49	468	0.67	461	0.56	458	1.74	465	0.04	462	0.23	443	18.75
United States	480	472	2.03	479	0.01	474	1.59	485	0.97	494	5.75	464	5.40	489	2.14
OECD average	500	500	0.00	499	0.01	499	0.03	500	0.00	500	0.00	500	0.01	500	0.00
Partners															
Argentina	566	563	0.32	565	0.05	558	2.29	572	1.14	571	0.55	579	4.17	576	2.43
Azerbaijan	611	609	0.27	607	0.73	600	7.34	620	2.87	603	3.30	624	7.67	613	0.17
Brazil	592	577	6.07	572	14.13	583	2.86	604	3.83	590	0.13	601	3.48	603	4.06
Bulgaria	522	520	0.09	526	0.30	517	0.66	519	0.24	521	0.04	528	0.76	542	6.19
Chile	591	579	4.53	585	1.00	598	1.54	605	5.67	596	0.72	595	0.38	600	1.74
Chinese Taipei	534	531	0.40	522	5.37	533	0.01	541	1.42	527	2.16	530	0.38	524	3.86
Colombia	642	626	3.36	637	0.46	638	0.39	630	1.86	653	2.55	664	4.95	633	1.20
Croatia	536	527	4.39	539	0.30	541	1.08	538	0.11	535	0.03	539	0.41	548	7.50
Estonia	503	503	0.02	502	0.00	500	0.38	497	1.43	498	1.17	508	1.31	503	0.00
Hong Kong-China	535	530	1.16	530	0.38	522	5.79	543	2.17	523	4.43	530	0.86	535	0.01
Indonesia	608	607	0.01	591	7.57	582	33.15	622	4.96	606	0.10	619	7.44	618	4.08
Israel	510	515	0.79	516	0.87	508	0.14	505	0.40	513	0.15	491	5.76	512	0.08
Jordan	608	603	1.35	606	0.13	606	0.19	602	1.44	611	0.36	604	0.89	626	9.51
Kyrgyzstan	581	584	0.69	572	3.66	562	11.64	578	0.45	589	3.09	597	13.82	581	0.00
Latvia	503	494	3.42	499	0.73	498	1.68	499	0.87	500	0.50	501	0.41	510	1.75
Liechtenstein	505	534	3.03	512	0.12	534	3.38	499	0.08	484	1.10	490	0.45	515	0.24
Lithuania	544	544	0.00	544	0.01	541	0.30	540	1.02	542	0.22	549	0.91	552	1.78
Macao-China	524	525	0.03	507	8.88	511	5.11	540	5.49	517	1.08	531	1.16	521	0.23
Montenegro	561	561	0.00	579	9.15	559	0.07	549	3.89	571	3.63	573	5.05	561	0.00
Qatar	566	556	3.92	561	0.91	565	0.02	560	0.85	576	5.87	562	0.62	582	8.65
Romania	591	573	9.96	581	3.27	588	0.25	590	0.01	599	1.47	599	0.77	604	5.36
Russian Federation	541	541	0.00	541	0.00	534	3.89	545	0.68	542	0.07	531	4.56	547	1.53
Serbia	524	515	1.98	520	0.43	536	8.14	529	1.12	522	0.08	529	1.17	528	1.03
Slovenia	505	508	0.14	501	0.72	518	3.82	497	1.73	500	1.11	507	0.15	514	2.90
Thailand	641	633	3.59	624	18.42	628	6.16	642	0.07	648	1.60	657	8.44	657	7.79
Tunisia	589	567	20.24	596	2.08	590	0.00	579	4.33	598	2.23	610	14.98	599	3.37
Uruguay	567	557	3.23	558	2.85	561	1.07	571	0.44	569	0.10	568	0.03	569	0.20



Table 12.17 [Part 2/2]
Variance in interest booklet means

	Boolet 8		Boolet 9		Boolet 10		Boolet 11		Boolet 12		Boolet 13		Chi-sq (df=12)
	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	
OECD													
Australia	476	8.72	469	0.45	460	2.56	472	2.54	464	0.28	464	0.47	35.3
Austria	505	0.08	504	0.20	513	1.71	501	0.71	507	0.07	499	1.55	12.0
Belgium	503	0.40	500	0.12	493	2.76	497	1.30	492	5.23	506	1.19	21.7
Canada	472	0.36	470	0.01	468	0.11	478	3.27	477	3.48	458	7.91	31.7
Czech Republic	488	0.07	484	0.27	491	0.42	482	0.67	492	0.85	485	0.10	20.3
Denmark	468	0.73	465	0.14	478	8.16	470	1.46	457	1.24	451	5.13	38.5
Finland	437	3.07	454	0.75	450	0.01	439	3.36	454	0.99	448	0.05	19.5
France	519	0.01	507	6.37	527	1.48	526	1.11	519	0.04	509	2.72	27.5
Germany	508	0.61	510	0.22	510	0.16	512	0.00	520	3.02	502	3.63	15.5
Greece	531	8.44	544	0.54	546	0.20	550	0.06	548	0.03	566	9.81	34.9
Hungary	525	0.40	527	0.99	537	8.58	509	6.11	511	4.15	516	1.71	29.9
Iceland	464	0.19	476	2.68	444	8.43	467	0.00	461	0.88	482	5.17	35.0
Ireland	479	0.13	492	4.58	478	0.29	479	0.20	472	2.72	482	0.04	12.6
Italy	535	2.31	526	0.79	534	2.67	527	0.40	534	2.26	530	0.10	32.2
Japan	509	0.20	512	0.01	515	0.25	513	0.03	523	3.71	514	0.10	36.9
Korea	491	0.68	491	1.57	494	1.60	486	0.00	491	1.03	481	0.95	15.8
Luxembourg	520	1.10	519	0.86	508	1.25	516	0.05	510	0.54	512	0.34	15.7
Mexico	615	1.08	607	0.64	614	0.50	623	5.68	609	0.13	612	0.05	36.2
Netherlands	449	0.01	443	1.70	442	2.44	458	1.84	459	3.08	453	0.40	48.9
New Zealand	453	1.88	454	1.61	461	0.00	461	0.00	470	2.19	457	0.44	19.4
Norway	469	0.50	474	0.04	465	1.72	473	0.00	464	2.43	457	7.23	43.1
Poland	512	5.88	497	0.41	514	7.52	497	0.43	505	1.11	503	0.37	33.1
Portugal	574	0.53	567	0.57	578	2.48	569	0.08	574	0.39	574	0.47	17.8
Slovak Republic	528	1.17	518	0.67	522	0.02	520	0.16	517	0.52	519	0.24	6.2
Spain	532	0.23	545	6.75	538	0.58	539	1.35	534	0.02	529	2.03	23.1
Sweden	458	0.32	452	0.55	467	5.09	441	4.90	452	0.41	467	4.72	25.7
Switzerland	508	1.61	499	0.95	500	0.87	500	0.56	496	4.19	508	0.87	22.2
Turkey	520	11.06	536	0.74	534	1.43	533	2.49	541	0.01	545	0.26	56.1
United Kingdom	475	5.87	462	0.20	475	6.17	475	4.18	464	0.00	447	11.40	53.3
United States	470	2.51	471	2.46	473	0.96	498	7.63	477	0.24	489	3.50	35.2
OECD average	500	0.00	499	0.01	501	0.05	500	0.01	500	0.00	499	0.03	0.2
Partners													
Argentina	556	1.59	554	7.42	570	0.31	565	0.06	576	2.19	569	0.14	22.7
Azerbaijan	607	0.91	602	4.14	621	4.87	604	1.86	639	32.82	603	2.88	69.8
Brazil	590	0.14	577	11.86	601	3.09	605	7.47	601	3.22	594	0.25	60.6
Bulgaria	519	0.17	514	1.27	510	3.40	522	0.00	511	2.36	544	9.72	25.2
Chile	590	0.06	575	5.09	589	0.17	594	0.19	584	1.01	591	0.01	22.1
Chinese Taipei	546	7.81	534	0.00	552	15.05	529	1.00	535	0.05	528	1.02	38.5
Colombia	661	3.25	622	9.87	664	7.30	660	6.90	635	0.88	640	0.07	43.0
Croatia	526	2.78	538	0.17	540	0.68	526	4.03	526	2.63	536	0.00	24.1
Estonia	507	1.25	502	0.04	511	4.61	488	7.97	500	0.27	510	2.55	21.0
Hong Kong-China	544	2.08	532	0.60	548	5.41	547	2.32	544	1.70	540	0.62	27.5
Indonesia	626	15.25	581	40.08	603	0.73	622	11.70	604	0.58	621	8.49	134.1
Israel	520	2.49	508	0.06	493	4.03	518	1.77	505	0.77	511	0.03	17.3
Jordan	618	3.66	597	5.08	604	0.59	610	0.26	599	5.38	633	22.98	51.8
Kyrgyzstan	590	4.91	567	10.99	595	7.99	584	0.31	572	2.55	572	3.01	63.1
Latvia	509	2.31	500	0.43	510	1.71	505	0.10	500	0.63	522	13.65	28.2
Liechtenstein	500	0.08	494	0.62	492	0.74	497	0.19	482	1.93	520	0.99	12.9
Lithuania	548	0.94	533	4.52	548	0.40	546	0.10	545	0.00	545	0.01	10.2
Macao-China	531	2.76	522	0.08	539	6.99	515	1.38	525	0.04	524	0.00	33.2
Montenegro	567	0.69	566	0.66	556	0.74	539	14.51	552	2.39	557	0.47	41.3
Qatar	568	0.23	549	11.26	554	3.81	573	2.19	558	1.98	585	14.48	54.8
Romania	596	0.63	583	1.44	605	5.55	591	0.00	579	3.64	601	2.29	34.6
Russian Federation	548	2.87	540	0.02	544	0.33	537	0.77	540	0.07	546	1.46	16.3
Serbia	508	7.15	518	1.15	526	0.13	515	2.72	518	0.77	529	1.09	26.9
Slovenia	503	0.14	511	1.26	500	0.63	485	7.22	504	0.09	512	1.60	21.5
Thailand	648	2.84	630	5.99	656	12.86	645	0.70	628	6.39	650	2.82	77.7
Tunisia	607	10.27	577	7.67	586	0.29	597	1.87	578	4.94	590	0.01	72.3
Uruguay	566	0.03	565	0.07	568	0.02	578	3.60	572	0.69	571	0.57	12.9



Table 12.18 [Part 2/2]
Variance in support booklet means

	Boolet 8		Boolet 9		Boolet 10		Boolet 11		Boolet 12		Boolet 13		Chi-sq (df=12)
	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	Mean	Z ²	
OECD													
Australia	489	0.08	488	0.01	473	13.15	495	3.52	479	4.23	479	4.74	61.1
Austria	528	2.85	519	0.23	523	0.83	531	3.77	527	2.11	495	7.39	54.8
Belgium	484	5.85	491	0.13	475	16.80	489	0.84	492	0.02	497	0.94	53.3
Canada	494	2.68	497	0.69	485	10.91	520	15.65	499	0.11	497	0.48	55.3
Czech Republic	468	16.72	497	6.48	488	0.07	483	0.35	502	7.04	493	1.54	54.0
Denmark	481	0.08	499	7.19	484	0.09	489	1.55	483	0.01	489	1.60	29.0
Finland	490	3.04	462	8.38	483	0.98	493	6.79	463	6.89	466	8.43	76.3
France	493	5.20	511	0.65	493	5.79	512	0.70	499	1.61	521	5.30	42.1
Germany	532	2.00	527	0.60	543	10.73	525	0.22	534	2.66	505	7.00	49.9
Greece	529	0.56	527	0.84	541	1.48	522	3.77	543	3.24	539	1.19	31.8
Hungary	505	1.41	518	1.30	532	14.71	496	5.80	512	0.01	494	13.56	49.6
Iceland	496	0.42	508	5.23	480	2.48	488	0.16	491	0.00	505	3.55	32.5
Ireland	486	0.05	495	5.09	481	0.37	489	1.27	482	0.13	482	0.26	24.9
Italy	502	5.49	514	0.44	505	2.71	492	25.01	510	0.20	520	6.21	94.6
Japan	492	14.55	473	0.32	466	0.41	471	0.08	441	14.55	457	3.82	93.2
Korea	490	0.92	481	6.55	507	3.69	492	0.36	504	1.77	507	5.07	58.7
Luxembourg	531	2.00	528	0.85	530	1.00	526	0.31	513	1.31	522	0.01	42.7
Mexico	522	9.82	535	0.01	547	6.33	513	16.42	559	36.93	550	11.85	345.7
Netherlands	447	0.02	452	0.85	447	0.05	452	0.69	448	0.00	429	15.73	68.9
New Zealand	466	0.54	465	0.87	467	0.38	454	9.62	467	0.24	472	0.16	52.4
Norway	494	1.40	488	0.28	482	0.17	499	3.66	466	7.83	489	0.30	46.2
Poland	512	0.02	513	0.00	535	13.10	482	40.38	535	16.80	531	18.99	162.7
Portugal	547	3.00	541	0.46	549	3.21	518	15.36	543	0.88	551	7.22	50.3
Slovak Republic	504	2.63	500	0.21	503	1.37	505	3.04	514	12.10	490	1.88	58.6
Spain	529	0.00	533	0.59	530	0.03	535	1.55	527	0.18	531	0.06	10.5
Sweden	485	4.56	468	0.32	473	0.08	488	8.53	469	0.09	474	0.20	31.2
Switzerland	510	0.02	508	0.24	513	0.22	510	0.00	510	0.02	506	0.99	17.0
Turkey	543	7.11	527	21.04	543	5.51	558	0.34	568	0.25	590	12.24	130.0
United Kingdom	472	0.32	464	1.72	475	1.92	475	1.12	461	4.95	453	15.69	46.3
United States	485	0.85	482	1.93	490	0.01	496	0.91	478	4.58	493	0.21	18.7
OECD average	500	0.00	500	0.00	501	0.03	500	0.00	501	0.00	501	0.01	0.2
Partners													
Argentina	478	10.15	513	1.23	519	3.00	498	1.81	529	11.66	493	4.56	82.2
Azerbaijan	506	33.45	511	29.68	555	6.23	507	16.15	587	61.76	563	15.10	343.3
Brazil	504	6.90	511	2.05	526	1.83	503	12.24	548	25.59	501	13.95	179.5
Bulgaria	516	2.78	522	0.57	528	0.00	515	2.57	537	1.19	537	1.57	39.6
Chile	564	0.04	559	0.59	564	0.03	560	0.74	535	12.25	551	4.73	56.2
Chinese Taipei	566	18.13	533	6.49	548	0.51	551	1.48	555	4.05	558	7.10	165.5
Colombia	524	9.79	543	0.14	542	0.21	520	12.01	585	32.08	547	0.07	125.2
Croatia	507	1.59	513	0.04	522	2.49	499	11.50	522	1.91	522	4.76	80.3
Estonia	513	10.99	489	2.39	496	0.03	492	0.72	507	3.25	513	9.88	75.0
Hong Kong-China	533	0.23	506	21.87	526	0.39	519	2.33	505	12.40	567	38.95	145.0
Indonesia	508	2.22	501	14.09	535	9.87	500	14.45	541	17.49	530	7.51	238.6
Israel	518	0.69	492	6.06	505	0.52	496	5.65	483	16.97	525	3.60	53.9
Jordan	553	0.05	553	0.09	561	1.05	512	45.01	573	9.10	560	0.61	182.7
Kyrgyzstan	469	25.66	474	25.40	528	15.89	488	4.38	507	1.09	517	6.67	259.7
Latvia	491	0.15	493	0.01	499	0.95	489	0.96	498	0.56	506	4.47	25.9
Liechtenstein	511	0.41	524	0.01	514	0.36	545	0.92	515	0.15	527	0.00	8.6
Lithuania	550	2.64	533	1.51	567	24.51	542	0.20	569	31.25	541	0.05	177.4
Macao-China	520	0.18	499	15.89	540	8.27	501	9.41	513	2.33	543	16.98	131.6
Montenegro	545	5.00	520	2.84	521	1.95	509	11.12	514	5.16	524	0.40	73.8
Qatar	493	12.62	508	2.85	514	0.45	474	50.01	509	1.62	516	0.27	185.3
Romania	529	1.40	531	0.95	564	8.58	532	0.36	541	0.14	540	0.06	46.1
Russian Federation	508	0.26	504	0.20	522	6.70	503	0.39	526	11.78	511	1.01	120.7
Serbia	528	1.34	515	1.20	523	0.17	501	6.74	522	0.02	524	0.26	27.0
Slovenia	503	0.00	505	0.06	497	1.33	486	4.68	509	0.90	506	0.25	66.1
Thailand	549	9.39	556	3.52	594	17.64	540	13.96	596	32.74	578	3.92	413.1
Tunisia	513	6.79	537	0.18	539	0.53	486	40.39	554	8.86	558	11.41	131.9
Uruguay	515	0.98	504	0.90	509	0.03	496	6.72	524	6.44	521	4.43	83.9



Booklets with the domain at the end of the booklet (mathematics in booklets 3, 7 and 13 and reading in booklets 2 and 9) have the highest parameters. The booklet effects for reading are more extreme than last cycle, possibly because the items in the major domain (science) include more words than the items of the major domain of last cycle (mathematics).

After scaling the PISA 2006 data for each country separately, the booklet parameters were added to the students' achievement scores for mathematics, reading, science, interest and support and mean performance scores could be compared across countries and across booklets. Tables 12.14 to 12.18 present results of testing the variance in booklet means by country (UH booklet excluded). The table rows represent countries and the columns booklets, the cells contain the mean performance by booklet and the squared difference between the observed and expected mean, divided by the error variance by booklet. The expected mean is the average of the booklet means, each weighted by the reciprocal of their error variance. The sum of the squared differences divided by their error variance is chi-square distributed with $13-1=12$ degrees of freedom. Significant values are in bold.

Taking the square root of the squared difference between observed and expected mean, divided by the error variance gives a z-score and is an indication of the magnitude of the difference between observed booklet mean and expected booklet mean. Significantly easier than expected booklets are bold and italic, significantly harder booklets than expected are bold. Shaded columns are booklets without items in the domain.

There is no significant booklet effect at the OECD level, because the booklet corrections controlled for this effect. Therefore, the booklet effects within countries are relative to the effect at OECD level. A plausible explanation for high chi-squares across domains of most countries is fatigue or speediness (Mexico, Colombia, Kyrgyzstan, Tunisia and Uruguay). In these cases the booklet means deviate most from the expected mean if the items of that domain appear at the end of the booklet. For some other countries, the reason for their relative high chi-squares across domains is less obvious (Azerbaijan, Brazil and Qatar).

The vast majority of booklets means for domains that are not included in the booklet (shaded columns for mathematics and reading) do not significantly differ from the expected booklet means, which is to be expected using the deviation contrast codes for booklets in the conditioning model.

Overview of the PISA cognitive reporting scales

PISA 2006 is the third PISA assessment and as such it is the third occasion on which reading, mathematics and science literacy scores have been reported. A central aim of PISA is to monitor trends over time in indicators based upon reading, mathematics and science literacy. In this section we review the stability of the PISA scales over time, with a view to:

- Setting out the range of scales that have been prepared over the past three PISA assessments;
- Describing their special features and appropriate use; and,
- Asking recommendations regarding future design elements of PISA.

Table 12.19 provides a listing of the 19 distinct cognitive scales that have been produced as part of PISA 2000, 2003 and 2006.² For the purpose of this overview, the cognitive scales are classified into three types: PISA overall literacy scales, PISA literacy scales and special purpose scales. PISA overall literacy scales are the key reporting scales that have been established for each domain, when that domain has been the major domain. The PISA literacy scales are sub-components of PISA overall literacy scales that were provided when a domain was the major domain. The special purpose scales are additional scales that can be used as interim and trend scales prior to the establishment of the related PISA overall literacy scales.



Table 12.19
Summary of PISA cognitive reporting scales

Name	Established	2000	2003	2006	Comment
PISA literacy scale					
PISA reading	2000	✓	✓	✓	Trends can be reported between any of the three cycles, by country or by subgroups within countries
PISA mathematics	2003		✓	✓	Trends can be reported between 2003 and 2006, by country or by subgroups within countries
PISA science	2006			✓	Provides the basis for future trend analysis by country or by subgroups within country
PISA literacy scales					
Reading scales					
Retrieving information	2000	✓			
Interpreting texts	2000	✓			
Reflection and evaluation	2000	✓			
Mathematics scales					
Quantity	2003		✓		
Uncertainty	2003		✓		
Space & shape	2003	✓	✓		Established in 2003 and then applied to 2000 with a rescaling (no conditioning). Trends can be reported for countries, but are not optimal for subgroups within countries.
Change & relationships	2003	✓	✓		Established in 2003 and then applied to 2000 with a rescaling (no conditioning). Trends can be reported for countries, but are not optimal for subgroups within countries.
Science scales					
Explaining phenomena scientifically	2006			✓	
Identifying scientific Issues	2006			✓	
Using scientific evidence	2006			✓	
Physical systems	2006			✓	Limited conditioning implemented permitting unbiased estimation by country and by gender. Results for other subgroups are not optimal.
Earth & space systems	2006			✓	Limited conditioning implemented permitting unbiased estimation by country and by gender. Results for other subgroups are not optimal.
Living systems	2006			✓	Limited conditioning implemented permitting unbiased estimation by country and by gender. Results for other subgroups are not optimal.
Special purpose scales					
Interim mathematics	2000	✓			
Interim science	2000	✓	✓		
Science trend 2003-2006	2006		✓	✓	Uses items that were common to PISA 2003 and 2006



In the table each scale is named, the database upon which it was established is given, the datasets for which it is provided are indicated; and comments are made about the scale's appropriate use. In the text following, further details are provided on these scales.

PISA overall literacy scales

The primary PISA reporting scales are PISA reading, PISA mathematics and PISA science. These scales were established in the year in which the respective domain was the major domain, since in that year the framework for the domain was fully developed and the domain was comprehensively assessed. When the overall literacy scale is established the mean of the scale is set at 500 and the standard deviation is set at 100 (for the pooled, equally weighted OECD countries) – for example, 500 on the PISA mathematics scale is the mean achievement of assessed students in OECD countries in 2003.

The intention is that these overall literacy scales will stay in place until the specification of the domain is changed or updated.

PISA literacy scales

Across the three PISA assessments a total of 13 scales have been prepared and reported. In PISA 2000, three reading aspect-based scales were prepared; in PISA 2003, four mathematics content-based scales were prepared; and in 2006 a total of six science scales were prepared.³

The scales are typically prepared only in the year in which a domain is a major domain, since when a domain is a major domain there are sufficient items in each sub-area to support the reporting of the scales. The one exception to this general practice is mathematics, for which the *space and shape* and *change and relationships* scales were reported for the PISA 2000 data as well as the PISA 2003 data. These scales, which were established in 2003 when mathematics was the major domain, could be applied to the 2000 data because only these two areas of mathematics had been assessed in PISA 2000, and sufficient common items were available to support the scaling.

For the 2000 data the mathematics scales were prepared using a methodology that permits trend analysis at the national level (or at the level of adjudicated regions), but the scales are not optimal for analysis at the level of student sub-groups.⁴

For science in PISA 2006, two alternative sets of scales were prepared. The first was a set of three process-based scales and the second was a set of three content-based scales. It is important to note that these are alternative scalings that each rely on the same test items. As such, it is inappropriate to jointly analyse scales that are selected from the alternative scalings. For example, it would not be meaningful or defensible to correlate or otherwise compare performance on the “Physical systems” scale, with performance on the *using scientific evidence* scale. Furthermore the content-based scales can be analysed at the national level (or at the level of adjudicated regions), and can be analysed by gender, but they are not optimal for use at the level of any other student sub-groups, whereas the process-based scales are suitable in addition for sub-group analyses.⁵

The metric of all of the PISA scales is set so that scales within a domain can be compared to each other and with the matching overall PISA reporting scale.⁶

Special purpose scales

There are three special purpose scales.

An interim mathematics scale was established and reported in PISA 2000. This scale was prepared to provide an overall mathematics score, and it used all of the mathematics items that were included in the PISA 2000 assessment. This scale was discontinued in 2003 when mathematics was the major domain and the alternative and more comprehensive PISA overall mathematics literacy scale was established.



An interim science scale was established and reported in PISA 2000. This scale was prepared to provide an overall science score, and it used all of the science items that were included in the PISA 2000 assessment. The PISA 2003 science data were linked to this scale so that the PISA 2003 science results were also reported on this interim science scale. For PISA 2006 this scale was not provided since science was the major domain and the alternative and more comprehensive overall PISA science scale was established.

To allow comparisons between science outcomes in 2003 and 2006 a science Trend 2003-2006 scale was prepared. This scale is based upon the science items that are common to PISA 2003 and 2006 and can be used to examine trends (on those common items) between 2003 and 2006. The PISA 2003 abilities that are based on the common items can be analysed at the national level (or at the level of adjudicated regions), and can be analysed by gender, but they are not optimal for use at the level of any other student sub-groups. The PISA 2006 abilities, associated with the fully developed overall *PISA science* scale, can be analysed by national subgroups as well.

OBSERVATIONS CONCERNING THE CONSTRUCTION OF THE PISA OVERALL LITERACY SCALES

A number of the PISA scales have been established to permit trend analyses. A review of the various links available and necessary to establish these scales is given below. Table 12.20 illustrates the six linkages of the PISA domains that are examined and discussed below. Links (1) and (2) are for reading 2000 to 2003 and 2003 to 2006 respectively, links (3) and (4) are for mathematics 2000 to 2003 and 2003 to 2006 respectively, links (5) and (6) are for science 2000 to 2003 and 2003 to 2006 respectively.

Table 12.20 also indicates in which data collections the domain was a major domain and on which occasions it was a minor domain. As a consequence one can note that on two occasions the links are major to minor (links (1) and (4)), on two occasions they are minor to minor (links (2) and (5)), and on two occasions they are minor to major (links (3) and (6)).

When a proficiency area is assessed as a major domain there are two key characteristics that distinguish it from a minor domain. First the framework for the area is fully developed and elaborated. Second the framework is comprehensively assessed since more assessment time is allocated to the major domain than is allocated to each of the minor domains.

Table 12.20
Linkage types among PISA domains 2000-2006

	2000		2003		2006
Reading	Major	(1) →	Minor	(2) →	Minor
Mathematics	Minor	(3) →	Major	(4) →	Minor
Science	Minor	(5) →	Minor	(6) →	Major

Framework development

For PISA 2000 a full and comprehensive framework was developed for reading to guide the assessment of reading as a major domain. Less fully articulated frameworks were developed to support the assessment of mathematics and science as minor domains.⁷

For PISA 2003, the mathematics framework was updated and fully developed to support a comprehensive assessment of mathematics. The reading and science frameworks were retained largely as they had been for PISA 2000.⁸



The key changes to the mathematics framework between 2000 and 2003 were:

- Addition of a theoretical underpinning of the mathematics assessment, expanding the rationale for the PISA emphasis on using mathematical knowledge and skills to solve problems encountered in life;
- Restructuring and expansion of domain content: expansion from two broad content areas (overarching ideas) to four; removal of all reference to mathematics curricular strands as a separate content categorisation (instead, definitions of the overarching ideas were expanded to include mention of the kinds of school mathematics topics associated with each);
- A more elaborated rationale for the existing balance between realistic mathematics and more traditional context-free items, in line with the literacy for life notion underlying OECD/PISA assessments;
- A redeveloped discussion of the relevant mathematical processes: a clearer and much enhanced link between the process referred to as mathematisation, the underlying mathematical competencies, and the competency clusters; and a better operationalisation of the competency classes through a more detailed description of the underlying proficiency demands they place on students;
- Considerable elaboration through addition of examples, including items from previous test administrations.

Clearly, the framework change involving an effective doubling of the mathematical content base of the study was of such significance that trend measures would be very seriously affected. Hence, only scale links to 2000 were possible, and the new framework provided the first comprehensive basis for the calculation of future trend estimates.

For PISA 2006, science was the major domain so the science framework was updated and fully developed to support a comprehensive assessment of science. The reading framework was retained largely as it had been for PISA 2000, and the mathematics framework as it had been for PISA 2003.⁹ The key changes to the science framework between 2003 and 2006 as they relate to comparison in the science scales over time were:

- A clearer separation of knowledge about science as a form of human enquiry from knowledge of science, meaning knowledge of the natural world as articulated in the different scientific disciplines. In particular, PISA 2006 gives greater emphasis to knowledge about science as an aspect of science performance, through the addition of elements that underscore students' knowledge about the characteristic features of science and scientific endeavour; and
- The addition of new components on the relationship between science and technology.

Both of these changes carry the potential to disrupt links with the previous special purpose science scales: the interim science and trend science scales.

Testing time and item characteristics

In each of PISA 2000, 2003 and 2006 a total of 390 unique minutes of testing material was used.¹⁰ The distribution of the testing minutes is given in Table 12.21. When a domain is assessed as a major domain then more minutes are devoted to it than for minor domains. For example 270 minutes were assigned to reading material in PISA 2000 to allow full coverage of the framework. Similarly, PISA 2003 included 210 minutes of mathematics material and PISA 2006 included 210 minutes of science material. When a domain is assessed as a minor domain the assessment is far less comprehensive and does not provide an in-depth assessment of the full framework that is developed when a domain is a major domain.



Table 12.21
Number of unique item minutes for each domain for each PISA assessment

	Reading	Mathematics	Science	Total
2000	270	60	60	390
2003	60	210	60	330 ¹
2006	60	120	210	390

1. 60 minutes were devoted to Problem solving.

It is also important to recognise that given the PISA test design (see Chapter 2) the change of major domains over time means that the testing experience for the majority of students will be different in each cycle - it becomes dominated by the new major domain. For example, the design for PISA 2006 used 13 booklets. Eleven of them comprised at least 50% of science material. For four of these the other 50% comprised only mathematics material, four were completed with a mixture of reading and mathematics material, and for one booklet the other 50% comprised only reading material. Two booklets contained only science material.

The links in terms of numbers of items in common for successive pairs of assessments are shown in Table 12.22.

Table 12.22
Number of link items between successive PISA assessments

	Reading	Mathematics	Science
As Major Domain	129	84	108
Links 2000-2003	28	20	25
Link 2003-2006	28	48	22

Characteristics of each of the links

Reading 2000 to 2003

The PISA reading scale was established in 2000 on the basis of a fully developed and articulated framework and a comprehensive assessment of that framework. In PISA 2003 a subset of 28 of the 2000 reading items was selected and used. Equating procedures reported in OECD (2005) were then used to report the PISA 2003 data on the established PISA reading scale.

The trend results for the OECD countries that participated in both PISA 2000 and 2003 showed that of 32 countries, 10 had a significant decline in mean score and 5 had a significant rise in mean score (OECD 2004). This number of significant changes was regarded as somewhat surprising.

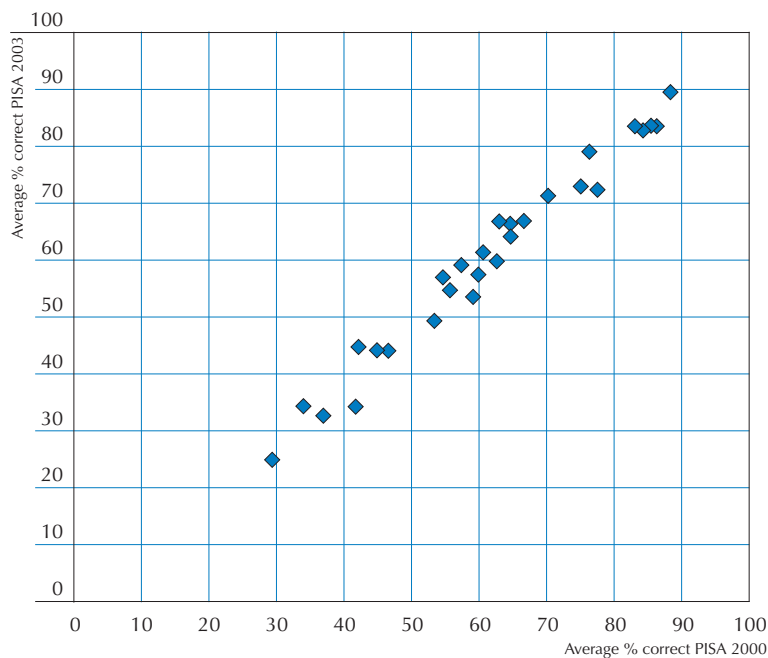
When reviewing the potential causes for this possible instability a number of potentially relevant issues were observed. First, there was a substantial test design change between PISA 2000 and 2003. The PISA 2003 design was fully balanced whereas the PISA 2000 design systematically placed minor domain items and some reading items at the end of the student booklets (see Adams & Wu, 2002). The complexity of the PISA 2000 design is such that the impact of this on the item parameter estimation and hence the equating is unclear. Second, the units that were selected from PISA 2000 for use in PISA 2003 were edited in minor ways. While none of the individual link items was edited, some items in the units were removed. As with the test design change, the impact of this change on the item parameter estimation and hence the equating is unclear. Third, the clusters of items that were used were not pre-existing clusters. In particular, units from PISA 2000 clusters one to seven were selected and reconstituted as two new clusters. Intact clusters of items could not be used from PISA 2000 since none of the individual pre-existing clusters provided an adequate coverage of the framework.



Table 12.23
Per cent correct for reading link items in PISA 2000 and PISA 2003

Item	% correct	
	2000	2003
R055Q01	84.4	82.9
R055Q02	53.4	49.1
R055Q03	62.7	59.8
R055Q05	77.7	72.5
R067Q01	88.5	89.7
R067Q04	54.7	57.0
R067Q05	62.9	67.1
R102Q04A	37.1	32.4
R102Q05	42.2	44.9
R102Q07	86.2	83.5
R104Q01	83.0	83.2
R104Q02	41.6	34.5
R104Q05	29.2	24.9
R111Q01	64.8	66.3
R111Q02B	34.2	34.0
R111Q06B	44.8	44.5
R219Q01	70.2	71.2
R219Q01E	57.4	59.3
R219Q02	76.5	78.8
R220Q01	46.8	44.4
R220Q02B	64.8	64.0
R220Q04	60.8	61.3
R220Q05	85.5	83.2
R220Q06	66.6	67.1
R227Q01	59.0	53.8
R227Q02	59.8	57.7
R227Q03	56.0	54.9
R227Q06	75.2	72.9

Figure 12.6
Scatter plot of per cent correct for reading link items
in PISA 2000 and PISA 2003





The percentage correct on reading items that link PISA 2000 and PISA 2003 are given in Table 12.23, with the corresponding scatterplot in Figure 12.6. To compute the percentage correct, all students were included from countries that were included in trend analysis between PISA 2000 and PISA 2003. For this analysis 25 OECD countries were included. Excluded were the United Kingdom, the Netherlands, Luxembourg, the Slovak Republic and Turkey.

The mean of the differences (PISA 2000 minus PISA 2003) is 1.11, and the standard deviation of the differences is 2.82.

Reading 2003 to 2006

To link the PISA 2006 data to the PISA reading scale the same items (units and clusters) as were used in PISA 2003 were again used. The trend results for the OECD countries that participated in both PISA 2003 and 2006 showed that of the 38 countries which could be compared, five had a significant decline in mean score and two had a significant rise in mean score (OECD 2007). The number of significant changes is less than reported for the 2000-2003 link.

A number of reasons might be conjectured as possible explanations of this lack of consistency. First, presenting a large number of reading items with a small number of mathematics and science items interspersed, provides for a very different test-taking experience for students compared to a test with a majority of mathematics items, and a few reading, general problem solving and science items interspersed.

Table 12.24
Per cent correct for reading link items in PISA 2003 and PISA 2006

Item	% correct	
	2003	2006
R055Q01	81.4	80.9
R055Q02	47.9	46.8
R055Q03	58.2	57.2
R055Q05	72.6	71.0
R067Q01	89.5	88.2
R067Q04	56.1	55.6
R067Q05	66.4	65.9
R102Q04A	32.4	32.2
R102Q05	43.1	42.8
R102Q07	81.8	82.9
R104Q01	83.0	80.3
R104Q02	34.3	32.9
R104Q05	25.3	22.8
R111Q01	64.9	63.4
R111Q02B	32.9	33.4
R111Q06B	43.3	40.9
R219Q01	69.6	68.4
R219Q01E	57.5	57.4
R219Q02	78.1	78.8
R220Q01	43.2	42.5
R220Q02B	63.5	61.2
R220Q04	62.1	59.2
R220Q05	83.2	81.0
R220Q06	67.1	66.4
R227Q01	53.7	52.3
R227Q02	57.9	55.0
R227Q03	54.4	53.3
R227Q06	71.3	69.3

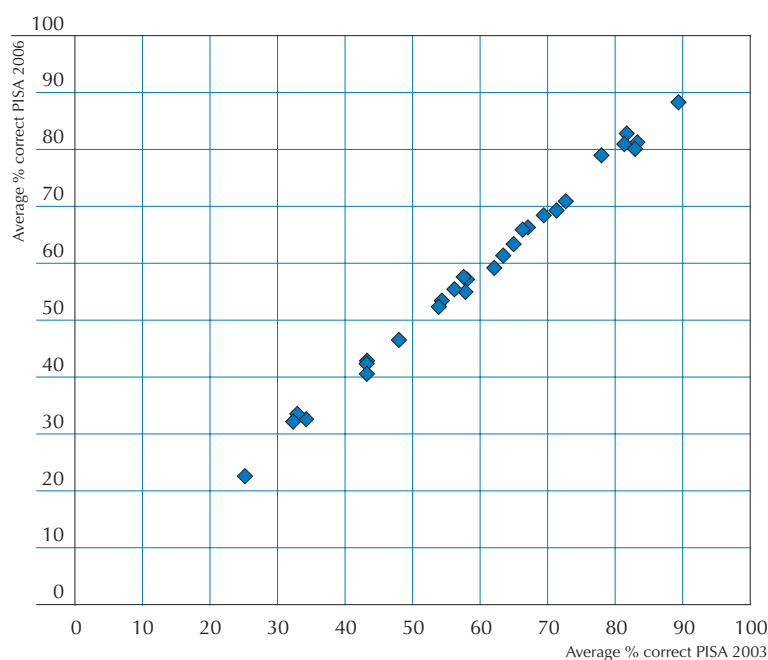


This may have impacted on the trend estimates. Second, the mix of reading items by aspect type was somewhat different between the two test administrations. In 2003 there was a larger proportion of score points in the reflection and evaluation aspect than had been the case for 2000.

The percentage correct on reading items that link PISA 2003 and PISA 2006 are given in Table 12.24, with the corresponding scatterplot in Figure 12.7. To compute the percentage correct, all students were included from countries that were included in these trend analyses. For percentage correct, 28 OECD countries were included. Excluded were the United Kingdom and the United States.

Figure 12.7

**Scatter plot of per cent correct for reading link items
in PISA 2003 and PISA 2006**



The mean of the differences (PISA 2003 minus PISA 2006) is 1.17, and the standard deviation of the differences is 1.07. The standard deviation of this difference is much less than that for 2003 to 2006 and most likely due to the use of identical items in identical clusters for the two assessments.

Mathematics 2000 to 2003

The mathematics framework that was prepared for PISA 2000 was preliminary and the assessment was restricted to two of the so-called big ideas – *space and shape*, and *change and relationships*. For the PISA 2003 assessment, when mathematics was a major domain, the framework was fully developed and the assessment was broadened to cover the four overarching ideas – *quantity*, *uncertainty*, *space and shape*, and *change and relationships*.

Given that the mathematics framework was fully developed for PISA 2003, the PISA mathematics scale was developed at that point. As PISA 2000 had covered two of the scales, two scale trend scales were developed that permit comparison of performance between 2000 and 2003 for *space and shape*, and *change and relationships*.



Mathematics 2003 to 2006

A selection of 48 mathematics items was selected from PISA 2003 and used again in PISA 2006.¹¹ Hence the change from 2003 to 2006 involved reducing the number of items by almost half, and as was the case when reading changed from major to minor domain, it was not possible to make such a reduction whilst retaining intact clusters. Four new clusters were formed for PISA 2006 from the units retained from PISA 2003. The trend results for the OECD countries that participated in both PISA 2003 and 2006 showed that of the 39 countries which could be compared four had a significant decline in mean score and four had a significant rise in mean score (OECD 2007). The magnitude and number of these changes is consistent with the figures for reading 2003 to 2006 and with figures observed in TIMSS.

Table 12.25
Per cent correct for mathematics link items in PISA 2003 and PISA 2006

Item	% correct	
	2003	2006
M033Q01	77.0	76.8
M034Q01	43.6	43.5
M155Q01	64.9	64.6
M155Q02	61.0	60.8
M155Q03	17.0	19.1
M155Q04	56.7	55.7
M192Q01	40.7	40.3
M273Q01	55.1	53.7
M302Q01	95.3	95.4
M302Q02	78.6	80.4
M302Q03	29.9	28.9
M305Q01	64.5	61.7
M406Q01	29.1	27.7
M406Q02	19.7	17.2
M408Q01	41.5	43.4
M411Q01	51.8	50.5
M411Q02	46.3	44.8
M420Q01	49.9	48.2
M421Q01	65.8	62.8
M421Q02	17.8	16.3
M421Q03	38.5	34.4
M423Q01	81.5	79.6
M442Q02	41.8	39.3
M446Q01	68.3	67.1
M446Q02	6.9	7.0
M447Q01	70.5	68.6
M462Q01	14.5	12.1
M464Q01	25.4	24.9
M474Q01	74.6	73.7
M496Q01	53.3	50.1
M496Q02	66.0	64.1
M559Q01	61.3	63.5
M564Q01	49.9	47.1
M564Q02	46.0	46.3
M571Q01	49.0	47.3
M598Q01	64.4	59.9
M603Q01	47.7	45.0
M603Q02	36.2	35.1
M710Q01	34.3	32.5
M800Q01	91.9	89.5
M803Q01	28.3	29.7
M810Q01	68.6	61.7
M810Q02	72.3	69.1
M810Q03	20.4	19.2
M828Q01	39.8	36.5
M828Q02	54.5	54.7
M828Q03	32.5	29.1
M833Q01	31.8	30.2

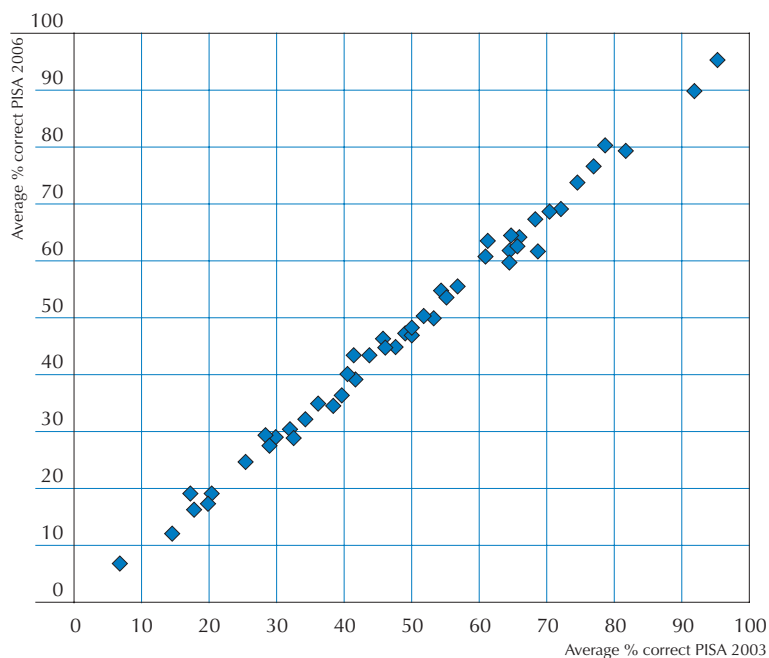


It is interesting to contrast these results with those observed for reading. At the item level the consistency seems somewhat less for mathematics than for reading, whereas at the scale level the consistency is comparable. It is our conjecture that the item-level inconsistency is caused because by the change from mathematics as a major domain to mathematics as a minor domain. Two specific aspects of the change are likely to have contributed to the observed degree of consistency. One is the fact that it was necessary to select a subset of items and form new trend clusters. The rearrangement of items into new clusters appears to have a small impact on relative item difficulty. The second is the fact that the items were presented to students in a different context from previously; specifically that the items were no longer from the dominant domain, rather they represented a smaller set of items presented amongst a much larger number of science items.

The percentage correct on mathematics items that link PISA 2003 and PISA 2006 are given in Table 12.25, with the corresponding scatterplot in Figure 12.8. To compute the percentage correct, all students were included from countries that were included in these trend analyses. For percentage correct, 29 OECD countries were included. The United Kingdom was excluded because it was excluded from PISA 2003.

Figure 12.8

Scatter plot of per cent correct for mathematics link items in PISA 2003 and PISA 2006



The mean of the differences (PISA 2003 minus PISA 2006) is 1.40, and the standard deviation of the differences is 1.77. This standard deviation is less than that for reading between 2000 and 2003 but greater than that for reading between 2003 and 2006. This is consistent with the fact that 2003 and 2006 designs were both balanced but, unlike the reading items, the mathematics link items between 2003 and 2006 were not presented in the same clusters.



Science 2000 to 2003

Science was a minor domain in both PISA 2000 and 2003. As such the assessment on both of these occasions was less comprehensive than it was in 2006, when a more fully articulated framework and more testing time was available. There were 25 items that were common to both PISA 2000 and 2003. The trend results for the OECD countries that participated in both PISA 2000 and 2003 showed that of 32 countries, 5 had a significant decline in mean score and 13 a significant rise in mean score (OECD 2004). This number of significant changes was regarded as somewhat surprising.

The number of inconsistencies between 2000 and 2003 was greater than expected at both the item-level and at the scale level. When reviewing the potential causes for this possible instability a number of potentially relevant issues were observed. First, as mentioned above for reading, there was a substantial test design change between PISA 2000 and 2003. The complexity of the PISA 2000 design is such that impact of this on the item parameter estimation and hence the equating is unclear. Second, the units that were selected from PISA 2000 for use in PISA 2003 were edited in minor ways. As with reading, while none of the link items was edited some items in the units were removed. And as with the test design change, the impact of this on the item parameter estimation and hence the equating is unclear. Third the clusters of items that were used were not pre-existing clusters. The material retained from the two PISA 2000 clusters was supplemented with a small number of new units, and reconstituted as two new clusters. Fourth, there were just 25 link items between these two assessments, and unlike mathematics these items were spread across all aspects of the framework. This number was less than desirable and was a result of choices made concerning the release of items following the 2000 assessment to illustrate the nature of the PISA assessment to the public.

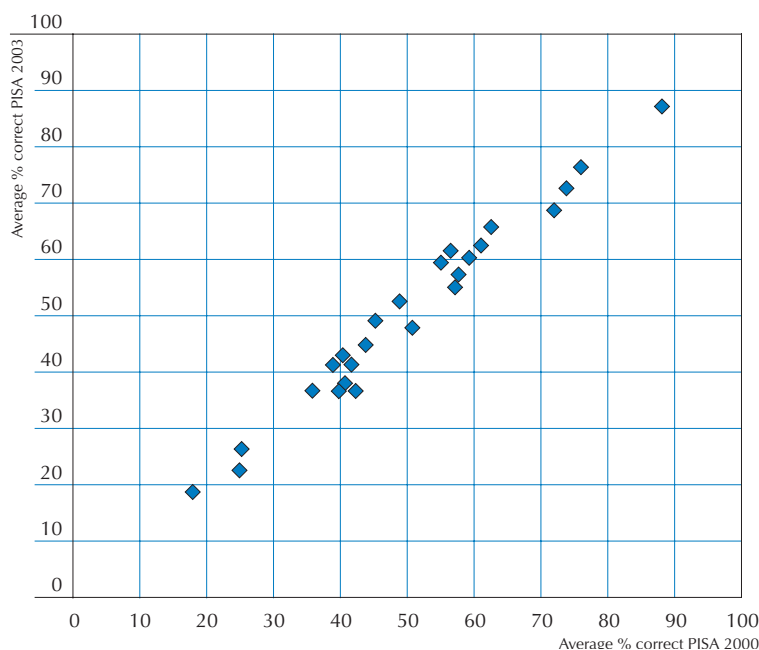
Table 12.26
Per cent correct for science link items in PISA 2000 and PISA 2003

Item	% correct	
	2000	2003
S114Q03	57.3	55.0
S114Q04	39.8	36.8
S114Q05	24.9	22.7
S128Q01	62.6	65.7
S128Q02	45.2	49.0
S128Q03	61.2	62.5
S129Q01	38.8	41.6
S129Q02	17.9	19.0
S131Q02	50.9	47.9
S131Q04	25.2	26.5
S133Q01	56.7	61.6
S133Q03	42.3	36.6
S133Q04	43.8	44.7
S213Q01	40.3	43.2
S213Q02	76.1	76.6
S252Q01	48.8	52.8
S252Q02	72.2	68.6
S252Q03	55.0	59.2
S256Q01	88.3	87.3
S268Q01	73.7	72.4
S268Q02	40.8	38.1
S268Q06	57.9	57.4
S269Q01	59.2	60.2
S269Q03	41.8	41.6
S269Q04	35.9	36.5



Figure 12.9

Scatter plot of per cent correct for science link items in PISA 2000 and PISA 2003



The percentage correct on science items that link PISA 2000 and PISA 2003 are given in Table 12.26, with the corresponding scatterplot in Figure 12.9. To compute the percentage correct, all students were included from countries that were included in these trend analyses. For percentage correct, 25 OECD countries were included. The United Kingdom, the Netherlands, Luxembourg, the Slovak Republic and Turkey were excluded because they did not participate in either PISA 2000 or PISA 2003 or were excluded for quality assurance reasons from one of PISA 2000 or PISA 2003.

The mean of the differences (PISA 2000 minus PISA 2003) is -0.28 , and the standard deviation of the differences is 2.79. This standard deviation is consistent with that observed for reading between 2000 and 2003.

Science 2003 to 2006

In PISA 2006, science was the major domain and as such it was comprehensively assessed on the basis of a newly developed and elaborated framework. As noted above there were quite substantial changes between the preliminary framework that had underpinned PISA 2000 and PISA 2003 assessments and the more fully developed framework used for PISA 2006. Note that in addition to the framework changes mentioned above there was an important change in the way science was assessed in PISA 2006, as compared with PISA 2003 and PISA 2000. First, to more clearly distinguish scientific literacy from reading literacy the PISA 2006 science test items required, on average, less reading than did the science items used in earlier PISA surveys. Second, as with each domain as it goes from a minor to a major domain the item pool (and therefore the testing experience for the majority of students) becomes dominated by the new major domain. For example, there were 108 science items used in PISA 2006, compared with 35 in PISA 2003; of these, just 22 items were common to PISA 2006 and PISA 2003 and 14 were common to PISA 2006 and PISA 2000.

So, as the first major assessment of science, the PISA 2006 assessment was used to establish the basis for the PISA science scale.

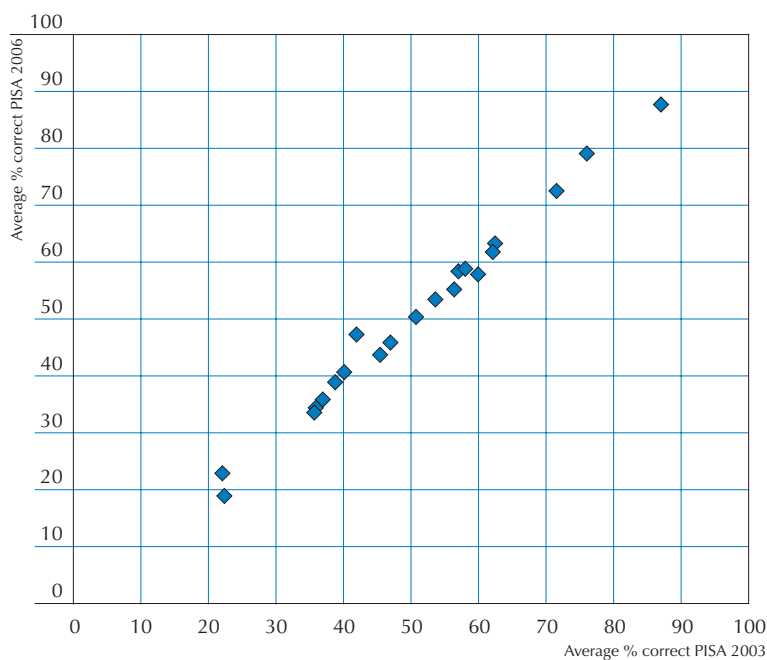


The percentage correct on science items that link PISA 2003 and PISA 2006 are given in Table 12.27, with the corresponding scatterplot and Figure 12.10. To compute the percentage correct, all students were included from countries that were included in these trend analyses. For percentage correct, 29 OECD countries were included. The United Kingdom was excluded because it was excluded from the PISA 2003 database.

Table 12.27
Per cent correct for science link items in PISA 2003 and PISA 2006

Item	% correct	
	2003	2006
S114Q03	53.6	53.6
S114Q04	35.9	34.4
S114Q05	22.4	18.8
S131Q02	46.9	46.2
S213Q01	41.9	47.4
S213Q02	76.2	79.2
S256Q01	87.0	87.5
S268Q01	71.7	72.5
S268Q02	36.9	36.1
S268Q06	56.6	55.4
S269Q01	60.0	57.9
S269Q03	40.1	40.7
S269Q04	35.6	33.8
S304Q01	45.5	43.8
S304Q02	62.0	62.1
S304Q03a	38.7	39.1
S304Q03b	50.7	50.6
S326Q01	58.2	58.7
S326Q02	62.6	63.4
S326Q03	57.2	58.3
S326Q04	22.2	22.8

Figure 12.10
Scatter plot of per cent correct for science link items
in PISA 2003 and PISA 2006





The mean of the differences (PISA 2000 minus PISA 2003) is -0.01 , and the standard deviation of the differences is 1.89 . This standard deviation is less than for science 2000-2003 but greater than that for reading 2003-2006. As with the previous observations regarding the standard deviations of the differences this is consistent with PISA test design changes.

For the purposes of trend analysis an additional trend scale has been established that is based upon those items that were common to both PISA 2003 and 2006. Details on the construction of this trend scale are given below and international results are provided in the initial report (OECD, 2007; p.369-370).

On the science trends scale that was produced from these 39 countries that participated in both PISA 2003 and PISA 2006, one had a significant decline in mean score and five had a significant rise in mean score (OECD 2007).

TRANSFORMING THE PLAUSIBLE VALUES TO PISA SCALES

Reading

The reading plausible values were equated to the PISA 2000 scale. Since the same items were used in PISA 2003 as in PISA 2006, and in each case the mean of the item parameter estimates is set at zero, the transformation was exactly the same as in PISA 2003.

For female students:

$$\text{PISA 2000 scale score} = ((0.8739 * \text{Logit} + 0.0970 - 0.5076) / 1.1002) * 100 + 500$$

For male students:

$$\text{PISA 2000 scale score} = ((0.8823 * \text{Logit} + 0.0204 - 0.5076) / 1.1002) * 100 + 500$$

For students with missing gender code:

$$\text{PISA 2000 scale score} = ((0.8830 * \text{Logit} + 0.0552 - 0.5076) / 1.1002) * 100 + 500$$

For details about equating procedures in 2003, the reader is referred to the *PISA 2003 Technical Report* (OECD, 2005).

Mathematics

For mathematics, the PISA 2006 plausible values were equated to the PISA 2003 scale. A shift of 0.0405 of a logit was required to align the 2003 and 2006 scales. After applying this shift, the same standardisation was used as in PISA 2003 (where -0.1344 is the OECD mean and 1.2838 the OECD standard deviation).

$$\text{PISA 2003 scale score} = (((\text{Logit} - 0.0405) + 0.1344) / 1.2838) * 100 + 500$$

Science

A new scale for science was established in PISA 2006. Therefore the only transformation to the plausible values was a standardisation to an OECD mean of 500 and OECD standard deviation of 100 (using an equally weighted, pooled database).

$$\text{PISA 2006 scale score} = ((\text{Logit} - 0.1797) / 1.0724) * 100 + 500$$

The same transformation parameters were used for the scales of science.

An additional set of plausible values was drawn for science link items only (in both 2003 and 2006) to provide estimates of trends in science. To equate the PISA 2006 abilities to the PISA 2003 scale, the following transformations was applied. After adding a shift that reflects the difference in mean item difficulty of the link



items in PISA 2003 and PISA 2006 (-0.1709), the same transformation was applied as in 2003. When the country means and the OECD average were computed, an upward trend was observed in most country means and the OECD average. To compensate for this shift, 13.0 PISA points were subtracted from the PISA 2006 country means to make the OECD average equal in both cycles (excluding the United Kingdom).

PISA 2003 scale score = $((1.0063 * (\text{Logit} - 0.1709) - 0.0155) + 0.0933) / 1.1085 * 100 + 500 - 13.0$.

Attitudinal scales

The interest and support attitudinal scales were established in PISA 2006 as well, so the same methodology as for science was applied.

For interest in science:

PISA 2006 scale score = $((\text{Logit} - 0.1785) / 1.1190) * 100 + 500$

For support of scientific enquiry

PISA 2006 scale score = $((\text{Logit} - 1.2694) / 0.8706) * 100 + 500$

LINK ERROR

Link errors estimated using the methodology discussed in Chapter 9 for the following five links; PISA mathematics scale 2003 to 2006, PISA reading scale 2000 to 2003, PISA reading scale 2000 to 2006, PISA reading scale 2003 to 2006, and science trend scale 2003 to 2006, are given in Table 12.28. Note that the value of 4.474 given for the PISA reading scale 2000 to 2003 link is a little larger than the value of 3.744, as reported in OECD (2005). Similarly for the Interim science scale the new estimate of 3.112 is a little larger than the previously reported value of 2.959. The differences in these values is due to the improved link error estimation method used for PISA 2006.

Table 12.28
Link error estimates

	Link Error on PISA Scale
Mathematics scale 2003 to 2006	1.382
Reading scale 2000 to 2003	5.307
Reading scale 2000 to 2006	4.976
Reading scale 2003 to 2006	4.474
Interim Science scale 2000 to 2003	3.112
Interim Science trend scale 2003 to 2006	4.963



Notes

1. Note that the USA was not included in the correlations with reading.
2. Note that this section refers to cognitive scales only. PISA has also produced a wide range of other scales that are affective or behavioural scales.
3. For a description of the content of the scales see the PISA framework publication (OECD, 2006, *Assessing Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006*).
4. This is because conditioning variables were not used in the construction of the scales for the PISA 2000 data.
5. This is because gender was the only conditioning variable used in the construction of the content-based scales.
6. Note, of course, that as mentioned above comparison across alternative scalings of the same domain are not appropriate.
7. The PISA 2000 frameworks were published as OECD (1999) *Measuring Student Knowledge and Skills: A new Framework for Assessment*.
8. The PISA 2003 frameworks were published as OECD (2003) *The PISA 2003 Assessment Framework: Mathematics, Reading, Science and Problem Solving Knowledge and Skills*.
9. The PISA 2006 frameworks were published as OECD (2006) *Assessment Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006*.
10. In 2003 the total testing time was also 390 minutes, but 60 minutes of that testing time was allocated to an assessment of Problem Solving skills.
11. Representing 120 minutes of testing time.



Reader's Guide

Country codes – the following country codes are used in this report:

OECD countries

AUS	Australia
AUT	Austria
BEL	Belgium
BEF	Belgium (French Community)
BEN	Belgium (Flemish Community)
CAN	Canada
CAE	Canada (English Community)
CAF	Canada (French Community)
CZE	Czech Republic
DNK	Denmark
FIN	Finland
FRA	France
DEU	Germany
GRC	Greece
HUN	Hungary
ISL	Iceland
IRL	Ireland
ITA	Italy
JPN	Japan
KOR	Korea
LUX	Luxembourg
LXF	Luxembourg (French Community)
LXG	Luxembourg (German Community)
MEX	Mexico
NLD	Netherlands
NZL	New Zealand
NOR	Norway
POL	Poland
PRT	Portugal
SVK	Slovak Republic
ESP	Spain
ESB	Spain (Basque Community)
ESC	Spain (Catalonian Community)
ESS	Spain (Castillian Community)
SWE	Sweden
CHE	Switzerland
CHF	Switzerland (French Community)
CHG	Switzerland (German Community)
CHI	Switzerland (Italian Community)

TUR	Turkey
GBR	United Kingdom
IRL	Ireland
SCO	Scotland
USA	United States

Partner countries and economies

ARG	Argentina
AZE	Azerbaijan
BGR	Bulgaria
BRA	Brazil
CHL	Chile
COL	Colombia
EST	Estonia
HKG	Hong Kong-China
HRV	Croatia
IDN	Indonesia
JOR	Jordan
KGZ	Kyrgyzstan
LIE	Liechtenstein
LTU	Lithuania
LVA	Latvia
LVL	Latvia (Latvian Community)
LVR	Latvia (Russian Community)
MAC	Macao-China
MNE	Montenegro
QAT	Qatar
ROU	Romania
RUS	Russian Federation
SRB	Serbia
SVN	Slovenia
TAP	Chinese Taipei
THA	Thailand
TUN	Tunisia
URY	Uruguay



References

- Adams, R.J., Wilson, M. & Wang, W.C.** (1997), The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement*, No. 21, pp. 1-23.
- Adams, R.J., Wilson, M. R. & Wu, M.L.** (1997), Multilevel item response models: An approach to errors in variables regression, *Journal of Educational and Behavioural Statistics*, No. 22 (1), pp. 46-75.
- Adams, R.J. & Wu, M.L.** (2002), *PISA 2000 Technical Report*, OECD, Paris.
- Bollen, K.A. & Long, S.J.** (1993) (eds.), *Testing Structural Equation Models*, Newbury Park: London.
- Beaton, A.E.** (1987), Implementing the new design: The NAEP 1983-84 technical report (Rep. No. 15-TR-20), Princeton, NJ: Educational Testing Service.
- Buchmann, C.** (2000), Family structure, parental perceptions and child labor in Kenya: What factors determine who is enrolled in school? *Soc. Forces*, No. 78, pp. 1349-79.
- Buchmann, C.** (2002), Measuring Family Background in International Studies of Education: Conceptual Issues and Methodological Challenges, in Porter, A.C. and Gamoran, A. (eds.). *Methodological Advances in Cross-National Surveys of Educational Achievement* (pp. 150-97), Washington, DC: National Academy Press.
- Creemers, B.P.M.** (1994), *The Effective Classroom*, London: Cassell.
- Cochran, W.G.** (1977), *Sampling techniques*, third edition, New York, NY: John Wiley and Sons.
- Ganzeboom, H.B.G., de Graaf, P.M. & Treiman, D.J.** (1992), A standard international socio-economic index of occupational status, *Social Science Research*, No. 21, pp. 1-56.
- Ganzeboom H.B. & Treiman, D.J.** (1996), Internationally comparable measures of occupational status for the 1988 international standard classification of occupations, *Social Science Research*, No. 25, pp. 201-239.
- Grisay, A.** (2003), Translation procedures in OECD/PISA 2000 international assessment, *Language Testing*, No. 20 (2), pp. 225-240.
- Hambleton, R.K., Swaminathan, H. & Rogers, H.J.** (1991), *Fundamentals of item response theory*, Newbury Park, London, New Delhi: SAGE Publications.
- Hambleton, R.K., Merenda, P.F. & Spielberger, C.D.** (2005), *Adapting Educational and Psychological Tests for Cross-Cultural Assessment*, IEA Lawrence Erlbaum Associates, Publishers, Mahwah, New Jersey.
- Harkness, J.A., Van de Vijver, F.J.R. & Mohler, P.Ph** (2003), *Cross-Cultural Survey Methods*, Wiley-Interscience, John Wiley & Sons, Inc., Hoboken, New Jersey.
- Harvey-Beavis, A.** (2002), Student and School Questionnaire Development, in R.J. Adams and M.L. Wu (eds.), *PISA 2000 Technical Report*, (pp. 33-38), OECD, Paris.
- International Labour Organisation (ILO)** (1990), *International Standard Classification of Occupations: ISCO-88*. Geneva: International Labour Office.
- Jöreskog, K.G. & Sörbom, Dag** (1993), *LISREL 8 User's Reference Guide*, Chicago: SSI.
- Judkins, D.R.** (1990), Fay's Method of Variance Estimation, *Journal of Official Statistics*, No. 6 (3), pp. 223-239.
- Kaplan, D.** (2000), *Structural equation modeling: Foundation and extensions*, Thousand Oaks: SAGE Publications.
- Keyfitz, N.** (1951), Sampling with probabilities proportionate to science: Adjustment for changes in probabilities, *Journal of the American Statistical Association*, No. 46, American Statistical Association, Alexandria, pp. 105-109.
- Kish, L.** (1992), Weighting for Unequal, *Pi. Journal of Official Statistics*, No. 8 (2), pp. 183-200.
- LISREL** (1993), K.G. Jöreskog & D. Sörbom, [computer software], Lincolnwood, IL: Scientific Software International, Inc.
- Lohr, S.L.** (1999), *Sampling: Design and Analysis*, Duxberry: Pacific Grove.
- Macaskill, G., Adams, R.J. & Wu, M.L.** (1998), Scaling methodology and procedures for the mathematics and science literacy, advanced mathematics and physics scale, in M. Martin and D.L. Kelly, Editors, *Third International Mathematics and Science Study, technical report Volume 3: Implementation and analysis*, Boston College, Chestnut Hill, MA.
- Masters, G.N. & Wright, B.D.** (1997), The Partial Credit Model, in W.J. van der Linden, & R.K. Hambleton (eds.), *Handbook of Modern Item Response Theory* (pp. 101-122), New York/Berlin/Heidelberg: Springer.

- Mislevy, R.J.** (1991), Randomization-based inference about latent variables from complex samples, *Psychometrika*, No. 56, pp. 177-196.
- Mislevy, R.J., Beaton, A., Kaplan, B.A. & Sheehan, K.** (1992), Estimating population characteristics from sparse matrix samples of item responses, *Journal of Educational Measurement*, No. 29 (2), pp. 133-161.
- Mislevy, R.J. & Sheehan, K.M.** (1987), Marginal estimation procedures, in Beaton, A.E., Editor, 1987. *The NAEP 1983-84 technical report*, National Assessment of Educational Progress, Educational Testing Service, Princeton, pp. 293-360.
- Mislevy, R.J. & Sheehan, K.M.** (1989), Information matrices in latent-variable models, *Journal of Educational Statistics*, No. 14, pp. 335-350.
- Mislevy, R.J. & Sheehan, K.M.** (1989), The role of collateral information about examinees in item parameter estimation, *Psychometrika*, No. 54, pp. 661-679.
- Monseur, C. & Berezner, A.** (2007), The Computation of Equating Errors in International Surveys in Education, *Journal of Applied Measurement*, No. 8 (3), 2007, pp. 323-335.
- Monseur, C.** (2005), An exploratory alternative approach for student non response weight adjustment, *Studies in Educational Evaluation*, No. 31 (2-3), pp. 129-144.
- Muthen, B. & L. Muthen** (1998), [computer software], *Mplus* Los Angeles, CA: Muthen & Muthen.
- Muthen, B., du Toit, S.H.C. & Spisic, D.** (1997), *Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes*, unpublished manuscript.
- OECD** (1999), *Classifying Educational Programmes. Manual for ISCED-97 Implementation in OECD Countries*, OECD, Paris.
- OECD** (2003), *Literacy Skills for the World of Tomorrow: Further results from PISA 2000*, OECD, Paris.
- OECD** (2004), *Learning for Tomorrow's World – First Results from PISA 2003*, OECD, Paris.
- OECD** (2005), *Technical Report for the OECD Programme for International Student Assessment 2003*, OECD, Paris.
- OECD** (2006), *Assessing Scientific, Reading and Mathematical Literacy: A framework for PISA 2006*, OECD, Paris.
- OECD** (2007), *PISA 2006: Science Competencies for Tomorrow's World*, OECD, Paris.
- PISA Consortium** (2006), *PISA 2006 Main Study Data Management Manual*, https://mypisa.acer.edu.au/images/mypisadoc/opmanual/pisa2006_data_management_manual.pdf
- Rasch, G.** (1960), Probabilistic models for some intelligence and attainment tests, Copenhagen: Nielsen & Lydiche.
- Routitski A. & Berezner, A.** (2006), Issues influencing the validity of cross-national comparisons of student performance. Data Entry Quality and Parameter Estimation. Paper presented at the Annual Meeting of the American Educational Research Association (AERA) in San Francisco, 7-11 April, https://mypisa.acer.edu.au/images/mypisadoc/area06routitsky_berezner.pdf
- Rust, K.** (1985), Variance Estimation for Complex Estimators in Sample Surveys, *Journal of Official Statistics*, No. 1, pp. 381-397.
- Rust, K.F. & Rao, J.N.K.** (1996), Variance Estimation for Complex Surveys Using Replication Techniques, *Survey Methods in Medical Research*, No. 5, pp. 283-310.
- Shao, J.** (1996), Resampling Methods in Sample Surveys (with Discussion), *Statistics*, No. 27, pp. 203-254.
- Särndal, C.-E., Swensson, B. & Wretman, J.** (1992), *Model Assisted Survey Sampling*, New York: Springer-Verlag.
- SAS® CALIS** (1992), W. Hartmann [computer software], Cary, NC: SAS Institute Inc.
- Scheerens, J.** (1990), School effectiveness and the development of process indicators of school functioning, *School effectiveness and school improvement*, No. 1, pp. 61-80.
- Scheerens, J. & Bosker, R.J.** (1997), *The Foundations of School Effectiveness*, Oxford: Pergamon.
- Schulz, W.** (2002), Constructing and Validating the Questionnaire composites, in R.J. Adams and M.L. Wu (eds.), *PISA 2000 Technical Report*, OECD, Paris.
- Schulz, W.** (2004), Mapping Student Scores to Item Responses, in W. Schulz and H. Sibberns (eds.), *IEA Civic Education Study, Technical Report* (pp. 127-132), Amsterdam: IEA.
- Schulz, W.** (2006a), *Testing Parameter Invariance for Questionnaire Indices using Confirmatory Factor Analysis and Item Response Theory*, Paper presented at the Annual Meetings of the American Educational Research Association (AERA) in San Francisco, 7-11 April.
- Schulz, W.** (2006b), *Measuring the socio-economic background of students and its effect on achievement in PISA 2000 and PISA 2003*, Paper presented at the Annual Meetings of the American Educational Research Association (AERA) in San Francisco, 7-11 April.
- Thorndike, R.L.** (1973), *Reading comprehension in fifteen countries*, New York, Wiley: and Stockholm: Almqvist & Wiksell.
- Travers, K.J. & Westbury, I.** (1989), *The IEA Study of Mathematics I: Analysis of Mathematics Curricula*, Oxford: Pergamon Press.



- Travers, K.J., Garden R.A. & Rosier, M.** (1989), Introduction to the Study, in Robitaille, D. A. and Garden, R. A. (eds), *The IEA Study of Mathematics II: Contexts and Outcomes of School Mathematics Curricula*, Oxford: Pergamon Press.
- Verhelst, N.** (2002), Coder and Marker Reliability Studies, in R.J. Adams & M.L. Wu (eds.), *PISA 2000 Technical Report*. OECD, Paris.
- Walberg, H.J.** (1984), Improving the productivity of American schools, *Educational Leadership*, No. 41, pp. 19-27.
- Walberg, H.** (1986), Synthesis of research on teaching, in M. Wittrock (ed.), *Handbook of research on teaching* (pp. 214-229), New York: Macmillan.
- Walker, M.** (2006), *The choice of Likert or dichotomous items to measure attitudes across culturally distinct countries in international comparative educational research*. Paper presented at the Annual Meetings of the American Educational Research Association (AERA) in San Francisco, 7-11 April.
- Walker, M.** (2007), Ameliorating Culturally-Based Extreme Response Tendencies To Attitude items, *Journal of Applied Measurement*, No. 8, pp. 267-278.
- Warm, T.A.** (1989), Weighted Likelihood Estimation of Ability in Item Response Theory, *Psychometrika*, No. 54 (3), pp. 427-450.
- Westat** (2007), *WesVar[®] 5.1* Computer software and manual, Rockville, MD: Author (also see <http://www.westat.com/wesvar>).
- Wilson, M.** (1994), Comparing Attitude Across Different Cultures: Two Quantitative Approaches to Construct Validity, in M. Wilson (ed.), *Objective measurement II: Theory into practice* (pp. 271-292), Norwood, NJ: Ablex.
- Wolter, K.M.** (2007), *Introduction to Variance Estimation*. Second edition, Springer: New York.
- Wu, M.L., Adams, R.J. & Wilson, M.R.** (1997), *ConQuest[®]: Multi-Aspect Test Software* [computer program manual], Camberwell, Vic.: Australian Council for Educational Research.



List of abbreviations – the following abbreviations are used in this report:

ACER	Australian Council for Educational Research	NPM	National Project Manager
AGFI	Adjusted Goodness-of-Fit Index	OECD	Organisation for Economic Cooperation and Development
BRR	Balanced Repeated Replication	PISA	Programme for International Student Assessment
CBAS	Computer Based Assessment of Science	PPS	Probability Proportional to Size
CFA	Confirmatory Factor Analysis	PGB	PISA Governing Board
CFI	Comparative Fit Index	PQM	PISA Quality Monitor
CITO	National Institute for Educational Measurement, The Netherlands	PSU	Primary Sampling Units
CIVED	Civic Education Study	QAS	Questionnaire Adaptations Spreadsheet
DIF	Differential Item Functioning	RMSEA	Root Mean Square Error of Approximation
ENR	Enrolment of 15-year-olds	RN	Random Number
ESCS	PISA Index of Economic, Social and Cultural Status	SC	School Co-ordinator
ETS	Educational Testing Service	SE	Standard Error
IAEP	International Assessment of Educational Progress	SD	Standard Deviation
I	Sampling Interval	SEM	Structural Equation Modelling
ICR	Inter-Country Coder Reliability Study	SMEG	Subject Matter Expert Group
ICT	Information Communication Technology	SPT	Study Programme Table
IEA	International Association for the Evaluation of Educational Achievement	TA	Test Administrator
INES	OECD Indicators of Education Systems	TAG	Technical Advisory Group
IRT	Item Response Theory	TCS	Target Cluster Size
ISCED	International Standard Classification of Education	TIMSS	Third International Mathematics and Science Study
ISCO	International Standard Classification of Occupations	TIMSS-R	Third International Mathematics and Science Study – Repeat
ISEI	International Socio-Economic Index	VENR	Enrolment for very small schools
MENR	Enrolment for moderately small school	WLE	Weighted Likelihood Estimates
MOS	Measure of size		
NCQM	National Centre Quality Monitor		
NDP	National Desired Population		
NEP	National Enrolled Population		
NFI	Normed Fit Index		
NIER	National Institute for Educational Research, Japan		
NNFI	Non-Normed Fit Index		



Table of contents

FOREWORD	3
CHAPTER 1 PROGRAMME FOR INTERNATIONAL STUDENT ASSESSMENT: AN OVERVIEW	19
Participation	21
Features of PISA	22
Managing and implementing PISA	23
Organisation of this report	23
READER'S GUIDE	25
CHAPTER 2 TEST DESIGN AND TEST DEVELOPMENT	27
Test scope and format	28
Test design	28
Test development centres	29
Development timeline	30
The PISA 2006 scientific literacy framework	30
Test development – cognitive items	31
▪ Item development process.....	31
▪ National item submissions.....	33
▪ National review of items.....	34
▪ International item review.....	35
▪ Preparation of dual (English and French) source versions.....	35
Test development – attitudinal items	35
Field trial	38
▪ Field trial selection.....	38
▪ Field trial design.....	39
▪ Despatch of field trial instruments.....	40
▪ Field trial coder training.....	40
▪ Field trial coder queries.....	40
▪ Field trial outcomes.....	41
▪ National review of field trial items.....	42
Main study	42
▪ Main study science items.....	42
▪ Main study reading items.....	44
▪ Main study mathematics items.....	45
▪ Despatch of main study instruments.....	46
▪ Main study coder training.....	46
▪ Main study coder query service.....	46
▪ Review of main study item analyses.....	47



CHAPTER 3 THE DEVELOPMENT OF THE PISA CONTEXT QUESTIONNAIRES	49
Overview	50
The conceptual structure	51
▪ A conceptual framework for PISA 2006	51
Research areas in PISA 2006	55
The development of the context questionnaires	57
The coverage of the questionnaire material	58
▪ Student questionnaire	58
▪ School questionnaire	59
▪ International options	59
▪ National questionnaire material	60
The implementation of the context questionnaires	60
 CHAPTER 4 SAMPLE DESIGN	 63
Target population and overview of the sampling design	64
Population coverage, and school and student participation rate standards	65
▪ Coverage of the PISA international target population	65
▪ Accuracy and precision	66
▪ School response rates	66
▪ Student response rates	68
Main study school sample	68
▪ Definition of the national target population	68
▪ The sampling frame	69
▪ Stratification	70
▪ Assigning a measure of size to each school	74
▪ School sample selection	74
▪ PISA and TIMSS or PIRLS overlap control	76
▪ Student samples	82
 CHAPTER 5 TRANSLATION AND CULTURAL APPROPRIATENESS OF THE TEST AND SURVEY MATERIAL	 85
Introduction	86
Development of source versions	86
Double translation from two source languages	87
PISA translation and adaptation guidelines	88
Translation training session	89
Testing languages and translation/adaptation procedures	89
International verification of the national versions	91
▪ VegaSuite	93
▪ Documentation	93
▪ Verification of test units	93
▪ Verification of the booklet shell	94
▪ Final optical check	94
▪ Verification of questionnaires and manuals	94
▪ Final check of coding guides	95
▪ Verification outcomes	95



Translation and verification outcomes – national version quality	96
▪ Analyses at the country level.....	96
▪ Analyses at the item level.....	103
▪ Summary of items lost at the national level, due to translation, printing or layout errors.....	104
CHAPTER 6 FIELD OPERATIONS	105
Overview of roles and responsibilities	106
▪ National project managers.....	106
▪ School coordinators.....	107
▪ Test administrators.....	107
▪ School associates.....	108
The selection of the school sample	108
Preparation of test booklets, questionnaires and manuals	108
The selection of the student sample	109
Packaging and shipping materials	110
Receipt of materials at the national centre after testing	110
Coding of the tests and questionnaires	111
▪ Preparing for coding.....	111
▪ Logistics prior to coding.....	113
▪ Single coding design.....	115
▪ Multiple coding.....	117
▪ Managing the process coding.....	118
▪ Cross-national coding.....	120
▪ Questionnaire coding.....	120
Data entry, data checking and file submission	120
▪ Data entry.....	120
▪ Data checking.....	120
▪ Data submission.....	121
▪ After data were submitted.....	121
The main study review	121
CHAPTER 7 QUALITY ASSURANCE	123
PISA quality control	124
▪ Comprehensive operational manuals.....	124
▪ National level implementation planning document.....	124
PISA quality monitoring	124
▪ Field trial and main study review.....	124
▪ Final optical check.....	126
▪ National centre quality monitor (NCQM) visits.....	126
▪ PISA quality monitor (PQM) visits.....	126
▪ Test administration.....	127
▪ Delivery.....	128
CHAPTER 8 SURVEY WEIGHTING AND THE CALCULATION OF SAMPLING VARIANCE	129
Survey weighting	130
The school base weight	131
▪ The school weight trimming factor.....	132

<ul style="list-style-type: none"> ▪ The student base weight 132 ▪ School non-response adjustment..... 132 ▪ Grade non-response adjustment..... 134 ▪ Student non-response adjustment..... 135 ▪ Trimming student weights..... 136 ▪ Comparing the PISA 2006 student non-response adjustment strategy with the strategy used for PISA 2003 136 ▪ The comparison..... 138 	
Calculating sampling variance	139
<ul style="list-style-type: none"> ▪ The balanced repeated replication variance estimator..... 139 ▪ Reflecting weighting adjustments..... 141 ▪ Formation of variance strata..... 141 ▪ Countries where all students were selected for PISA..... 141 	
CHAPTER 9 SCALING PISA COGNITIVE DATA	143
The mixed coefficients multinomial logit model	144
<ul style="list-style-type: none"> ▪ The population model..... 145 ▪ Combined model..... 146 	
Application to PISA	146
<ul style="list-style-type: none"> ▪ National calibrations..... 146 ▪ National reports..... 147 ▪ International calibration 153 ▪ Student score generation..... 153 	
Booklet effects	155
Analysis of data with plausible values	156
Developing common scales for the purposes of trends	157
<ul style="list-style-type: none"> ▪ Linking PISA 2003 and PISA 2006 for reading and mathematics 158 ▪ Uncertainty in the link..... 158 	
CHAPTER 10 DATA MANAGEMENT PROCEDURES	163
Introduction	164
KeyQuest	167
Data management at the national centre	167
<ul style="list-style-type: none"> ▪ National modifications to the database 167 ▪ Student sampling with <i>KeyQuest</i>..... 167 ▪ Data entry quality control 167 	
Data cleaning at ACER	171
<ul style="list-style-type: none"> ▪ Recoding of national adaptations..... 171 ▪ Data cleaning organisation..... 171 ▪ Cleaning reports..... 171 ▪ General recodings..... 171 	
Final review of the data	172
<ul style="list-style-type: none"> ▪ Review of the test and questionnaire data 172 ▪ Review of the sampling data 172 	
Next steps in preparing the international database	172



CHAPTER 11 SAMPLING OUTCOMES	175
Design effects and effective sample sizes	187
▪ Variability of the design effect.....	191
▪ Design effects in PISA for performance variables.....	191
Summary analyses of the design effect	203
▪ Countries with outlying standard errors.....	205
CHAPTER 12 SCALING OUTCOMES	207
International characteristics of the item pool	208
▪ Test targeting.....	208
▪ Test reliability.....	208
▪ Domain inter-correlations.....	208
▪ Science scales.....	215
Scaling outcomes	216
▪ National item deletions.....	216
▪ International scaling.....	219
▪ Generating student scale scores.....	219
Test length analysis	219
Booklet effects	221
▪ Overview of the PISA cognitive reporting scales.....	232
▪ PISA overall literacy scales.....	234
▪ PISA literacy scales.....	234
▪ Special purpose scales.....	234
Observations concerning the construction of the PISA overall literacy scales	235
▪ Framework development.....	235
▪ Testing time and item characteristics.....	236
▪ Characteristics of each of the links.....	237
Transforming the plausible values to PISA scales	246
▪ Reading.....	246
▪ Mathematics.....	246
▪ Science.....	246
▪ Attitudinal scales.....	247
Link error	247
CHAPTER 13 CODING AND MARKER RELIABILITY STUDIES	249
Homogeneity analyses	251
Multiple marking study outcomes (variance components)	254
▪ Generalisability coefficients.....	254
International coding review	261
▪ Background to changed procedures for PISA 2006.....	261
▪ ICR procedures.....	261
▪ Outcomes.....	264
▪ Cautions.....	270



CHAPTER 14 DATA ADJUDICATION	271
Introduction	272
▪ Implementing the standards – quality assurance	272
▪ Information available for adjudication	273
▪ Data adjudication process	273
General outcomes	274
▪ Overview of response rate issues	274
▪ Detailed country comments	275
 CHAPTER 15 PROFICIENCY SCALE CONSTRUCTION	 283
Introduction	284
Development of the described scales	285
▪ Stage 1: Identifying possible scales	285
▪ Stage 2: Assigning items to scales	286
▪ Stage 3: Skills audit	286
▪ Stage 4: Analysing field trial data	286
▪ Stage 5: Defining the dimensions	287
▪ Stage 6: Revising and refining with main study data	287
▪ Stage 7: Validating	287
Defining proficiency levels	287
Reporting the results for PISA science	290
▪ Building an item map	290
▪ Levels of scientific literacy	292
▪ Interpreting the scientific literacy levels	299
 CHAPTER 16 SCALING PROCEDURES AND CONSTRUCT VALIDATION OF CONTEXT QUESTIONNAIRE DATA	 303
Overview	304
Simple questionnaire indices	304
▪ Student questionnaire indices	304
▪ School questionnaire indices	307
▪ Parent questionnaire indices	309
Scaling methodology and construct validation	310
▪ Scaling procedures	310
▪ Construct validation	312
▪ Describing questionnaire scale indices	314
Questionnaire scale indices	315
▪ Student scale indices	315
▪ School questionnaire scale indices	340
▪ Parent questionnaire scale indices	342
▪ The PISA index of economic, social and cultural status (ESCS)	346
 CHAPTER 17 VALIDATION OF THE EMBEDDED ATTITUDINAL SCALES	 351
Introduction	352
International scalability	353
▪ Analysis of item dimensionality with exploratory and confirmatory factor analysis	353
▪ Fit to item response model	353



▪ Reliability.....	355
▪ Differential item functioning.....	355
▪ Summary of scalability.....	357
Relationship and comparisons with other variables.....	357
▪ Within-country student level correlations with achievement and selected background variables.....	358
▪ Relationships between embedded scales and questionnaire.....	360
▪ Country level correlations with achievement and selected background variables.....	361
▪ Variance decomposition.....	363
▪ Observations from other cross-national data collections.....	363
▪ Summary of relations with other variables.....	364
Conclusion.....	364
CHAPTER 18 INTERNATIONAL DATABASE.....	367
Files in the database.....	368
▪ Student files.....	368
▪ School file.....	370
▪ Parent file.....	370
Records in the database.....	371
▪ Records included in the database.....	371
▪ Records excluded from the database.....	371
Representing missing data.....	371
How are students and schools identified?.....	372
Further information.....	373
REFERENCES.....	375
APPENDICES.....	379
Appendix 1 PISA 2006 main study item pool characteristics.....	380
Appendix 2 Contrast coding used in conditioning.....	389
Appendix 3 Design effect tables.....	399
Appendix 4 Changes to core questionnaire items from 2003 to 2006.....	405
Appendix 5 Mapping of ISCED to years.....	411
Appendix 6 National household possession items.....	412
Appendix 7 Exploratory and confirmatory factor analyses for the embedded items.....	414
Appendix 8 PISA consortium, staff and consultants.....	416



LIST OF BOXES

Box 1.1	Core features of PISA 2006.....	22
---------	---------------------------------	----

LIST OF FIGURES

Figure 2.1	Main study Interest in Science item.....	36
Figure 2.2	Main study Support for Scientific Enquiry item.....	36
Figure 2.3	Field trial Match-the-opinion Responsibility item.....	37
Figure 3.1	Conceptual grid of variable types.....	52
Figure 3.2	The two-dimensional conceptual matrix with examples of variables collected or available from other sources.....	54
Figure 4.1	School response rate standard.....	67
Figure 6.1	Design for the single coding of science and mathematics.....	115
Figure 6.2	Design for the single coding of reading.....	116
Figure 9.1	Example of item statistics in Report 1.....	148
Figure 9.2	Example of item statistics in Report 2.....	149
Figure 9.3	Example of item statistics shown in Graph B.....	150
Figure 9.4	Example of item statistics shown in Graph C.....	151
Figure 9.5	Example of item statistics shown in Table D.....	151
Figure 9.6	Example of summary of dodgy items for a country in Report 3a.....	152
Figure 9.7	Example of summary of dodgy items in Report 3b.....	152
Figure 10.1	Data management in relation to other parts of PISA.....	164
Figure 10.2	Major data management stages in PISA.....	166
Figure 10.3	Validity reports - general hierarchy.....	170
Figure 11.1	Standard error on a mean estimate depending on the intraclass correlation.....	188
Figure 11.2	Relationship between the standard error for the science performance mean and the intraclass correlation within explicit strata (PISA 2006).....	205
Figure 12.1	Item plot for mathematics items.....	210
Figure 12.2	Item plot for reading items.....	211
Figure 12.3	Item plot for science items.....	212
Figure 12.4	Item plot for interest items.....	213
Figure 12.5	Item plot for support items.....	214
Figure 12.6	Scatter plot of per cent correct for reading link items in PISA 2000 and PISA 2003.....	238
Figure 12.7	Scatter plot of per cent correct for reading link items in PISA 2003 and PISA 2006.....	240
Figure 12.8	Scatter plot of per cent correct for mathematics link items in PISA 2003 and PISA 2006.....	242
Figure 12.9	Scatter plot of per cent correct for science link items in PISA 2000 and PISA 2003.....	244
Figure 12.10	Scatter plot of per cent correct for science link items in PISA 2003 and PISA 2006.....	245



Figure 13.1	Variability of the homogeneity indices for science items in field trial	250
Figure 13.2	Average of the homogeneity indices for science items in field trial and main study	251
Figure 13.3	Variability of the homogeneity indices for each science item in the main study	252
Figure 13.4	Variability of the homogeneity indices for each reading item in the main study	252
Figure 13.5	Variability of the homogeneity indices for each mathematics item	252
Figure 13.6	Variability of the homogeneity indices for the participating countries in the main study	253
Figure 13.7	Example of ICR report (reading)	269
<hr/>		
Figure 14.1	Attained school response rates	274
<hr/>		
Figure 15.1	The relationship between items and students on a proficiency scale	285
Figure 15.2	What it means to be at a level	289
Figure 15.3	A map for selected science items	291
Figure 15.4	Summary descriptions of the six proficiency levels on the science scale	294
Figure 15.5	Summary descriptions of six proficiency levels in <i>identifying scientific issues</i>	295
Figure 15.6	Summary descriptions of six proficiency levels in <i>explaining phenomena scientifically</i>	297
Figure 15.7	Summary descriptions of six proficiency levels in <i>using scientific evidence</i>	300
<hr/>		
Figure 16.1	Summed category probabilities for fictitious item	314
Figure 16.2	Fictitious example of an item map	315
Figure 16.3	Scatterplot of country means for ESCS 2003 and ESCS 2006	347
<hr/>		
Figure 17.1	Distribution of item fit mean square statistics for embedded attitude items	354
Figure 17.2	An example of the ESC plot for item S408RNA	356
Figure 17.3	Scatterplot of mean mathematics interest against mean mathematics for PISA 2003	363

LIST OF TABLES

Table 1.1	PISA 2006 participants	21
<hr/>		
Table 2.1	Cluster rotation design used to form test booklets for PISA 2006	29
Table 2.2	Test development timeline for PISA 2006	30
Table 2.3	Science field trial all items	39
Table 2.4	Allocation of item clusters to test booklets for field trial	39
Table 2.5	Science main study items (item format by competency)	43
Table 2.6	Science main study items (item format by knowledge type)	44
Table 2.7	Science main study items (knowledge category by competency)	44
Table 2.8	Reading main study items (item format by aspect)	44
Table 2.9	Reading main study items (item format by text format)	45
Table 2.10	Reading main study items (text type by aspect)	45
Table 2.11	Mathematics main study items (item format by competency cluster)	45
Table 2.12	Mathematics main study items (item format by content category)	46
Table 2.13	Mathematics main study items (content category by competency cluster)	46

Table 3.1	Themes and constructs/variables in PISA 2006.....	56
Table 4.1	Stratification variables	71
Table 4.2	Schedule of school sampling activities	78
Table 5.1	Countries sharing a common version with national adaptations	90
Table 5.2	PISA 2006 translation/adaptation procedures.....	91
Table 5.3	Mean deviation and root mean squared error of the item by country interactions for each version.....	97
Table 5.4	Correlation between national item parameter estimates for Arabic versions.....	99
Table 5.5	Correlation between national item parameter estimates for Chinese versions.....	99
Table 5.6	Correlation between national item parameter estimates for Dutch versions.....	99
Table 5.7	Correlation between national item parameter estimates for English versions.....	99
Table 5.8	Correlation between national item parameter estimates for French versions.....	99
Table 5.9	Correlation between national item parameter estimates for German versions.....	100
Table 5.10	Correlation between national item parameter estimates for Hungarian versions.....	100
Table 5.11	Correlation between national item parameter estimates for Italian versions.....	100
Table 5.12	Correlation between national item parameter estimates for Portuguese versions.....	100
Table 5.13	Correlation between national item parameter estimates for Russian versions.....	100
Table 5.14	Correlation between national item parameter estimates for Spanish versions	100
Table 5.15	Correlation between national item parameter estimates for Swedish versions	100
Table 5.16	Correlation between national item parameter estimates within countries.....	101
Table 5.17	Variance estimate.....	102
Table 5.18	Variance estimates	103
Table 6.1	Design for the multiple coding of science and mathematics.....	118
Table 6.2	Design for the multiple coding of reading.....	118
Table 8.1	Non-response classes	133
Table 9.1	Deviation contrast coding scheme	154
Table 10.1	Double entry discrepancies per country: field trial data.....	169
Table 11.1	Sampling and coverage rates.....	178
Table 11.2	School response rates before replacement.....	182
Table 11.3	School response rates after replacement.....	184
Table 11.4	Student response rates after replacement.....	185
Table 11.5	Standard errors for the PISA 2006 combined science scale	189
Table 11.6	Design effect 1 by country, by domain and cycle.....	193
Table 11.7	Effective sample size 1 by country, by domain and cycle.....	194
Table 11.8	Design effect 2 by country, by domain and cycle.....	195
Table 11.9	Effective sample size 2 by country, by domain and cycle.....	196
Table 11.10	Design effect 3 by country, by domain and by cycle.....	197



Table 11.11	Effective sample size 3 by country, by domain and cycle	198
Table 11.12	Design effect 4 by country, by domain and cycle.....	199
Table 11.13	Effective sample size 4 by country, by domain and cycle	200
Table 11.14	Design effect 5 by country, by domain and cycle.....	201
Table 11.15	Effective sample size 5 by country, by domain and cycle	202
Table 11.16	Median of the design effect 3 per cycle and per domain across the 35 countries that participated in every cycle.....	203
Table 11.17	Median of the standard errors of the student performance mean estimate for each domain and PISA cycle for the 35 countries that participated in every cycle	203
Table 11.18	Median of the number of participating schools for each domain and PISA cycle for the 35 countries that participated in every cycle.....	204
Table 11.19	Median of the school variance estimate for each domain and PISA cycle for the 35 countries that participated in every cycle.....	204
Table 11.20	Median of the intraclass correlation for each domain and PISA cycle for the 35 countries that participated in every cycle.....	204
Table 11.21	Median of the within explicit strata intraclass correlation for each domain and PISA cycle for the 35 countries that participated in every cycle	205
Table 11.22	Median of the percentages of school variances explained by explicit stratification variables, for each domain and PISA cycle for the 35 countries that participated in every cycle	205
<hr/>		
Table 12.1	Number of sampled student by country and booklet.....	209
Table 12.2	Reliabilities of each of the four overall scales when scaled separately.....	215
Table 12.3	Latent correlation between the five domains	215
Table 12.4	Latent correlation between science scales	215
Table 12.5	Items deleted at the national level	216
Table 12.6	Final reliability of the PISA scales	216
Table 12.7	National reliabilities for the main domains.....	217
Table 12.8	National reliabilities for the science subscales.....	218
Table 12.9	Average number of not-reached items and missing items by booklet.....	219
Table 12.10	Average number of not-reached items and missing items by country.....	220
Table 12.11	Distribution of not-reached items by booklet	221
Table 12.12	Estimated booklet effects on the PISA scale.....	221
Table 12.13	Estimated booklet effects in logits	221
Table 12.14	Variance in mathematics booklet means	222
Table 12.15	Variance in reading booklet means.....	224
Table 12.16	Variance in science booklet means.....	226
Table 12.17	Variance in interest booklet means	228
Table 12.18	Variance in support booklet means.....	230
Table 12.19	Summary of PISA cognitive reporting scales	233
Table 12.20	Linkage types among PISA domains 2000-2006	235
Table 12.21	Number of unique item minutes for each domain for each PISA assessments.....	237
Table 12.22	Numbers of link items between successive PISA assessments.....	237
Table 12.23	Per cent correct for reading link items in PISA 2000 and PISA 2003	238
Table 12.24	Per cent correct for reading link items in PISA 2003 and PISA 2006	239
Table 12.25	Per cent correct for mathematics link items in PISA 2003 and PISA 2006	241

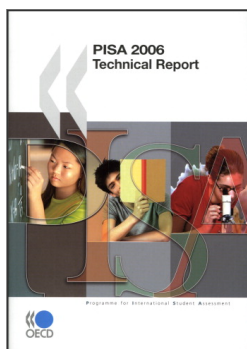
Table 12.26	Per cent correct for science link items in PISA 2000 and PISA 2003	243
Table 12.27	Per cent correct for science link items in PISA 2003 and PISA 2006	245
Table 12.28	Link error estimates	247
<hr/>		
Table 13.1	Variance components for mathematics.....	255
Table 13.2	Variance components for science	256
Table 13.3	Variance components for reading.....	257
Table 13.4	Generalisability estimates for mathematics.....	258
Table 13.5	Generalisability estimates for science	259
Table 13.6	Generalisability estimates for reading	260
Table 13.7	Examples of flagged cases	263
Table 13.8	Count of analysis groups showing potential bias, by domain.....	264
Table 13.9	Comparison of codes assigned by verifier and adjudicator	265
Table 13.10	Outcomes of ICR analysis part 1	265
Table 13.11	ICR outcomes by country and domain	266
<hr/>		
Table 15.1	Scientific literacy performance band definitions on the PISA scale	293
<hr/>		
Table 16.1	ISCO major group white-collar/blue-collar classification	306
Table 16.2	ISCO occupation categories classified as science-related occupations	307
Table 16.3	OECD means and standard deviations of WL estimates	311
Table 16.4	Median, minimum and maximum percentages of between-school variance for student-level indices across countries.....	313
Table 16.5	Household possessions and home background indices.....	316
Table 16.6	Scale reliabilities for home possession indices in OECD countries	317
Table 16.7	Scale reliabilities for home possession indices in partner countries/economies	318
Table 16.8	Item parameters for interest in science learning (INTSCIE).....	318
Table 16.9	Item parameters for enjoyment of science (JOYSCIE)	319
Table 16.10	Model fit and estimated latent correlations for interest in and enjoyment of science learning.....	319
Table 16.11	Scale reliabilities for interest in and enjoyment of science learning.....	320
Table 16.12	Item parameters for instrumental motivation to learn science (INSTSCIE).....	320
Table 16.13	Item parameters for future-oriented science motivation (SCIEFUT).....	321
Table 16.14	Model fit and estimated latent correlations for motivation to learn science	321
Table 16.15	Scale reliabilities for instrumental and future-oriented science motivation.....	322
Table 16.16	Item parameters for science self-efficacy (SCIEEFF).....	322
Table 16.17	Item parameters for science self-concept (SCSCIE).....	323
Table 16.18	Model fit and estimated latent correlations for science self-efficacy and science self-concept.....	323
Table 16.19	Scale reliabilities for science self-efficacy and science self-concept.....	324
Table 16.20	Item parameters for general value of science (GENSCIE).....	324
Table 16.21	Item parameters for personal value of science (PERSCIE).....	325
Table 16.22	Model fit and estimated latent correlations for general and personal value of science.....	325
Table 16.23	Scale reliabilities for general and personal value of science.....	326
Table 16.24	Item parameters for science activities (SCIEACT)	326



Table 16.25	Scale reliabilities for the science activities index	327
Table 16.26	Item parameters for awareness of environmental issues (ENVAWARE)	327
Table 16.27	Item parameters for perception of environmental issues (ENVPERC)	328
Table 16.28	Item parameters for environmental optimism (ENVOPT)	328
Table 16.29	Item parameters for responsibility for sustainable development (RESPDEV)	328
Table 16.30	Model fit environment-related constructs	329
Table 16.31	Estimated latent correlations for environment-related constructs	329
Table 16.32	Scale reliabilities for environment-related scales in OECD countries	330
Table 16.33	Scale reliabilities for environment-related scales in non-OECD countries	330
Table 16.34	Item parameters for school preparation for science career (CARPREP)	331
Table 16.35	Item parameters for student information on science careers (CARINFO)	331
Table 16.36	Model fit and estimated latent correlations for science career preparation indices	332
Table 16.37	Scale reliabilities for science career preparation indices	332
Table 16.38	Item parameters for science teaching: interaction (SCINTACT)	333
Table 16.39	Item parameters for science teaching: hands-on activities (SCHANDS)	333
Table 16.40	Item parameters for science teaching: student investigations (SCINVEST)	333
Table 16.41	Item parameters for science teaching: focus on models or applications (SCAPPLY)	334
Table 16.42	Model fit for CFA with science teaching and learning	334
Table 16.43	Estimated latent correlations for constructs related to science teaching and learning	335
Table 16.44	Scale reliabilities for scales to science teaching and learning in OECD countries	336
Table 16.45	Scale reliabilities for scales to science teaching and learning in partner countries/economies	336
Table 16.46	Item parameters for ICT Internet/entertainment use (INTUSE)	337
Table 16.47	Item parameters for ICT program/software use (PRGUSE)	337
Table 16.48	Item parameters for ICT self-confidence in Internet tasks (INTCONF)	337
Table 16.49	Item parameters for ICT self-confidence in high-level ICT tasks (HIGHCONF)	338
Table 16.50	Model fit for CFA with ICT familiarity items	338
Table 16.51	Estimated latent correlations for constructs related to ICT familiarity	339
Table 16.52	Scale reliabilities for ICT familiarity scales	339
Table 16.53	Item parameters for teacher shortage (TCSHORT)	340
Table 16.54	Item parameters for quality of educational resources (SCMATEDU)	340
Table 16.55	Item parameters for school activities to promote the learning of science (SCIPROM)	341
Table 16.56	Item parameters for school activities for learning environmental topics (ENVLEARN)	341
Table 16.57	Scale reliabilities for school-level scales in OECD countries	341
Table 16.58	Scale reliabilities for environment-related scales in partner countries/economies	342
Table 16.59	Item parameters for science activities at age 10 (PQSCIACT)	343
Table 16.60	Item parameters for parent's perception of school quality (PQSCHOOL)	343
Table 16.61	Item parameters for parent's views on importance of science (PQSCIMP)	343
Table 16.62	Item parameters for parent's reports on science career motivation (PQSCCAR)	344
Table 16.63	Item parameters for parent's view on general value of science (PQGENSCI)	344
Table 16.64	Item parameters for parent's view on personal value of science (PQPERSCI)	344
Table 16.65	Item parameters for parent's perception of environmental issues (PQENPERC)	345
Table 16.66	Item parameters for parent's environmental optimism (PQENVOPT)	345



Table 16.67	Scale reliabilities for parent questionnaire scales.....	345
Table 16.68	Factor loadings and internal consistency of ESCS 2006 in OECD countries.....	347
Table 16.69	Factor loadings and internal consistency of ESCS 2006 in partner countries/economies.....	348
<hr/>		
Table 17.1	Student-level latent correlations between mathematics, reading, science, embedded interest and embedded support.....	354
Table 17.2	Summary of the IRT scaling results across countries.....	355
Table 17.3	Gender DIF table for embedded attitude items.....	357
Table 17.4	Correlation amongst attitudinal scales, performance scales and HISEI.....	358
Table 17.5	Correlations for science scale.....	359
Table 17.6	Loadings of the achievement, interest and support variables on three varimax rotated components.....	360
Table 17.7	Correlation between embedded attitude scales and questionnaire attitude scales.....	361
Table 17.8	Rank order correlation five test domains, questionnaire attitude scales and HISEI.....	362
Table 17.9	Intra-class correlation (rho).....	362
<hr/>		
Table A1.1	2006 Main study reading item classification.....	380
Table A1.2	2006 Main study mathematics item classification.....	381
Table A1.3	2006 Main study science item classification (cognitive).....	383
Table A1.4	2006 Main study science embedded item classification (interest in learning science topics).....	387
Table A1.5	2006 Main study science embedded item classification (support for scientific enquiry).....	388
<hr/>		
Table A2.1	2006 Main study contrast coding used in conditioning for the student questionnaire variables.....	389
Table A2.2	2006 Main study contrast coding used in conditioning for the ICT questionnaire variables.....	396
Table A2.3	2006 Main study contrast coding used in conditioning for the parent questionnaire variables and other variables.....	397
<hr/>		
Table A3.1	Standard errors of the student performance mean estimate by country, by domain and cycle.....	399
Table A3.2	Sample sizes by country and cycle.....	400
Table A3.3	School variance estimate by country, by domain and cycle.....	401
Table A3.4	Intraclass correlation by country, by domain and cycle.....	402
Table A3.5	Within explicit strata intraclass correlation by country, by domain and cycle.....	403
Table A3.6	Percentages of school variance explained by explicit stratification variables, by domain and cycle.....	404
<hr/>		
Table A4.1	Student questionnaire.....	405
Table A4.2	ICT familiarity questionnaire.....	407
Table A4.3	School questionnaire.....	408
<hr/>		
Table A5.1	Mapping of ISCED to accumulated years of education.....	411
<hr/>		
Table A6.1	National household possession items.....	412
<hr/>		
Table A7.1	Exploratory and confirmatory factor analyses (EFA and CFA) for the embedded items.....	414



From:
PISA 2006 Technical Report

Access the complete publication at:
<https://doi.org/10.1787/9789264048096-en>

Please cite this chapter as:

OECD (2009), "Scaling Outcomes", in *PISA 2006 Technical Report*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/9789264048096-13-en>

This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of OECD member countries.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to rights@oecd.org. Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at info@copyright.com or the Centre français d'exploitation du droit de copie (CFC) at contact@cfcopies.com.