Appendix 10B: TIMSS 2019 Trend Scoring Reliability for Human Scored Items

TIMSS 2019 Trend Scoring Reliability for Human Scored Items—Grade 4 Mathematics

Country	Score Point Agreement			Diagnostic Code Agreement		
	Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement Across Items		Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement Across Items	
		Minimum	Maximum		Minimum	Maximum
Australia	98	90	100	97	89	100
Bahrain	99	96	100	98	91	100
Belgium (Flemish)	96	87	99	96	86	99
Bulgaria	99	95	100	99	95	100
Canada	95	76	99	94	75	99
Chile	96	85	100	96	85	100
Chinese Taipei	98	93	100	97	92	100
Croatia	98	90	100	97	90	100
Cyprus	97	87	100	97	87	100
Czech Republic	97	89	100	97	89	100
Denmark	96	84	100	94	78	100
England	98	86	100	97	86	100
Finland	99	93	100	98	93	100
France	97	87	100	96	87	100
Georgia	95	67	100	95	67	100
Germany	98	88	100	97	88	100
Hong Kong SAR	96	81	100	96	81	100
Hungary	97	93	99	97	92	99
Iran, Islamic Rep. of	98	92	99	97	92	99
Ireland	98	85	100	97	85	100
Italy	97	87	100	97	86	100
Japan	97	89	100	97	89	100
Korea, Rep. of	99	94	100	99	94	100
Lithuania	98	89	100	97	84	100
Netherlands	98	84	100	97	84	100
New Zealand	97	86	100	96	86	100
Northern Ireland	98	87	100	98	87	100

TIMSS 2019 Trend Scoring Reliability for Human Scored Items—Grade 4 Mathematics (continued)

Country	Score Point Agreement			Diagnostic Code Agreement		
	Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement Across Items		Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement Across Items	
		Minimum	Maximum		Minimum	Maximum
Norway (5)	98	91	100	97	90	100
Oman	95	87	100	93	72	100
Poland	97	89	100	97	89	100
Portugal	98	88	100	98	87	100
Qatar	97	89	100	95	77	100
Saudi Arabia	93	85	98	93	84	98
Serbia	96	79	100	95	79	100
Singapore	99	95	100	99	94	100
Slovak Republic	97	82	100	96	81	100
South Africa (5)	97	94	99	97	93	99
Spain	96	87	100	96	86	100
Sweden	96	78	100	96	77	100
United Arab Emirates	97	89	100	96	86	100
United States	97	85	100	97	85	100
International Average	**97**	**87**	**100**	**96**	**86**	**100**

TIMSS 2019 Trend Scoring Reliability for Human Scored Items—Grade 4 Science

Country	Score Point Agreement			Diagnostic Code Agreement		
	Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement Across Items		Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement Across Items	
		Minimum	Maximum		Minimum	Maximum
Australia	93	85	99	93	85	99
Bahrain	91	82	98	89	77	98
Belgium (Flemish)	92	82	99	91	82	99
Bulgaria	96	85	100	95	85	100
Canada	91	79	99	90	79	99
Chile	87	73	98	87	73	98
Chinese Taipei	93	81	100	93	76	100
Croatia	92	73	99	91	73	99
Cyprus	91	77	99	91	77	99
Czech Republic	91	69	98	90	69	98
Denmark	86	75	97	86	75	97
England	89	70	99	88	70	99
Finland	93	84	100	93	84	100
France	93	73	99	92	73	99
Georgia	88	69	97	87	69	95
Germany	93	82	99	92	82	99
Hong Kong SAR	89	82	96	88	80	96
Hungary	93	86	99	92	83	99
Iran, Islamic Rep. of	93	84	99	92	84	99
Ireland	91	74	99	91	74	99
Italy	93	81	100	93	81	100
Japan	91	83	99	90	83	99
Korea, Rep. of	95	88	100	95	88	100
Lithuania	93	80	99	93	80	99
Netherlands	94	78	99	93	78	99
New Zealand	90	85	97	90	81	97
Northern Ireland	92	82	99	91	82	99
Norway (5)	87	67	99	86	67	99
Oman	83	62	97	82	62	97
Poland	92	76	98	92	76	98
Portugal	95	87	98	95	87	98
Qatar	86	71	98	85	71	98

TIMSS 2019 Trend Scoring Reliability for Human Scored Items—Grade 4 Science (continued)

Country	Score Point Agreement			Diagnostic Code Agreement		
	Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement Across Items		Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement Across Items	
		Minimum	Maximum		Minimum	Maximum
Saudi Arabia	81	67	95	80	67	95
Serbia	90	81	99	89	77	99
Singapore	95	90	100	95	90	100
Slovak Republic	91	73	98	91	73	98
Spain	86	66	100	86	66	100
Sweden	90	77	99	90	77	99
United Arab Emirates	91	80	99	90	80	99
United States	94	83	100	93	83	100
International Average	**91**	**78**	**99**	**90**	**77**	**99**

TIMSS 2019 Trend Scoring Reliability for Human Scored Items—Grade 8 Mathematics

Country	Score Point Agreement			Diagnostic Code Agreement		
	Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement Across Items		Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement Across Items	
		Minimum	Maximum		Minimum	Maximum
Australia	96	89	100	96	89	100
Bahrain	98	89	100	95	65	100
Canada *	88	74	93	85	74	93
Chile	95	86	100	94	78	100
Chinese Taipei	97	90	100	96	79	100
England	95	63	100	95	63	100
Georgia	94	79	100	91	52	100
Hong Kong SAR	92	69	100	92	63	100
Hungary	97	83	100	95	74	100
Iran, Islamic Rep. of	97	89	100	97	89	100
Ireland	97	84	100	97	84	100
Israel	97	89	100	96	89	100
Italy	97	89	100	97	89	100
Japan	96	68	100	94	68	100
Jordan	99	97	100	98	93	100
Korea, Rep. of	94	49	100	93	49	100
Lithuania	98	92	100	96	74	100
Malaysia	96	84	100	93	60	100
New Zealand	95	81	100	95	81	100
Norway (9)	94	71	100	93	71	100
Oman	96	83	99	94	68	99
Qatar	96	83	100	95	80	100
Russian Federation	97	84	100	96	81	100
Saudi Arabia	97	88	100	97	88	99
Singapore	97	78	100	97	78	100
South Africa (9)	97	83	100	92	27	100
Sweden	96	74	100	95	74	100
Turkey	96	75	100	94	75	100
United Arab Emirates	97	86	100	96	85	100
United States	97	84	100	97	84	100
International Average	**96**	**81**	**100**	**95**	**74**	**100**

* Canada participated in trend scoring reliability for the benchmarking participants Ontario and Quebec.

TIMSS 2019 Trend Scoring Reliability for Human Scored Items—Grade 8 Science

Country	Score Point Agreement			Diagnostic Code Agreement		
	Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement Across Items		Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement Across Items	
		Minimum	Maximum		Minimum	Maximum
Australia	92	78	97	91	78	97
Bahrain	89	74	98	87	65	98
Canada *	90	80	98	89	70	98
Chile	87	70	99	86	66	99
Chinese Taipei	92	81	99	90	74	99
England	91	81	98	90	81	98
Georgia	90	75	99	88	66	99
Hong Kong SAR	90	77	99	89	77	99
Hungary	92	83	97	91	78	97
Iran, Islamic Rep. of	93	81	100	92	81	100
Ireland	93	84	100	92	84	100
Israel	92	85	100	91	81	100
Italy	92	83	100	91	73	100
Japan	92	81	100	91	75	100
Jordan	98	94	100	98	94	100
Korea, Rep. of	95	87	99	95	83	99
Lithuania	95	78	100	94	78	100
Malaysia	90	79	98	89	79	98
New Zealand	91	82	98	90	82	98
Norway (9)	91	82	99	90	79	99
Oman	86	71	97	84	65	97
Qatar	89	80	99	87	73	99
Russian Federation	86	66	98	85	65	98
Saudi Arabia	88	78	99	87	72	99
Singapore	94	82	100	94	82	100
South Africa (9)	95	90	100	94	88	100
Sweden	91	78	100	91	78	100
Turkey	90	77	97	88	66	97
United Arab Emirates	92	88	98	91	84	98
United States	93	80	98	92	76	98
International Average	**91**	**80**	**99**	**90**	**76**	**99**

* Canada participated in trend scoring reliability for the benchmarking participants Ontario and Quebec.