

Appendix 17B: Within-Country Scoring Reliability for Human Scored Items in the TIMSS 2019 Problem Solving and Inquiry Tasks

Within-Country Scoring Reliability for Human Scored Items in the TIMSS 2019 Problem Solving and Inquiry Tasks—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
Austria	98	96	99	96	96	97
Canada	93	88	98	93	88	97
Chile	97	97	98	96	96	97
Chinese Taipei	95	93	97	94	93	95
Croatia	98	98	99	98	98	98
Czech Republic	96	93	99	95	93	97
Denmark	97	95	98	96	95	98
England	97	97	97	96	95	97
Finland	100	100	100	100	100	100
France	99	99	99	98	97	99
Georgia	96	94	97	96	94	97
Germany	96	95	98	94	93	95
Hong Kong SAR	100	100	100	100	100	100
Hungary	96	95	96	96	95	96
Italy	98	98	99	98	98	98
Korea, Rep. of	98	97	98	98	97	98
Lithuania	97	95	100	97	95	98
Malta	93	91	95	90	88	91
Netherlands	92	88	96	91	88	93
Norway (5)	97	97	98	97	97	97
Portugal	96	94	97	94	94	94
Qatar	97	96	97	97	96	97
Russian Federation	97	95	98	97	95	98
Singapore	99	99	100	99	99	99
Slovak Republic	97	96	99	96	96	97
Spain	97	94	99	97	94	99
Sweden	99	98	99	99	98	99
Turkey (5)	99	99	99	99	98	99
United Arab Emirates	99	98	99	98	98	98
United States	98	96	99	97	96	98
International Average	97	96	98	96	95	97
Benchmarking Participants						
Ontario, Canada	94	88	100	94	88	100
Quebec, Canada	95	93	98	94	93	96
Moscow City, Russian Fed.	99	98	99	99	98	99
Madrid, Spain	97	95	99	96	95	97
Abu Dhabi, UAE	100	100	100	98	97	100
Dubai, UAE	98	98	99	98	97	98

Within-Country Scoring Reliability for Human Scored Items in the TIMSS 2019 Problem Solving and Inquiry Tasks—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
Austria	95	93	98	95	93	97
Canada	89	81	97	88	81	96
Chile	94	91	97	93	91	96
Chinese Taipei	94	88	100	93	88	99
Croatia	92	85	97	92	85	97
Czech Republic	91	86	97	90	84	96
Denmark	93	85	98	93	85	98
England	93	88	98	93	88	97
Finland	97	93	100	97	92	99
France	93	89	97	93	89	97
Georgia	91	79	98	91	76	98
Germany	94	88	97	94	82	97
Hong Kong SAR	100	100	100	100	100	100
Hungary	89	79	98	89	79	97
Italy	97	96	99	97	95	99
Korea, Rep. of	97	94	100	97	94	100
Lithuania	94	89	98	93	88	97
Malta	91	85	99	91	85	98
Netherlands	89	83	96	89	82	94
Norway (5)	91	80	96	90	80	95
Portugal	94	90	98	94	90	98
Qatar	95	91	100	94	91	98
Russian Federation	94	90	99	94	90	98
Singapore	95	89	100	94	86	100
Slovak Republic	96	93	99	96	93	99
Spain	91	86	98	90	86	97
Sweden	90	86	96	89	84	96
Turkey (5)	95	90	99	95	90	98
United Arab Emirates	93	91	97	93	91	96
United States	96	93	99	96	93	99
International Average	93	88	98	93	88	98
Benchmarking Participants						
Ontario, Canada	89	81	100	88	81	97
Quebec, Canada	90	79	100	89	79	100
Moscow City, Russian Fed.	97	92	99	96	92	99
Madrid, Spain	91	84	99	90	84	97
Abu Dhabi, UAE	93	88	99	92	88	96
Dubai, UAE	91	84	97	90	84	96

Within-Country Scoring Reliability for Human Scored Items in the TIMSS 2019 Problem Solving and Inquiry Tasks—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
Chile	97	94	99	97	94	99
Chinese Taipei	98	94	100	98	94	100
England	98	96	99	98	95	99
Finland	99	99	100	99	99	100
France	97	93	99	97	93	99
Georgia	95	89	99	95	89	99
Hong Kong SAR	100	100	100	100	100	100
Hungary	96	93	99	96	93	99
Israel	98	96	99	98	96	99
Italy	97	93	99	97	93	99
Korea, Rep. of	98	96	99	98	96	99
Lithuania	97	92	100	97	92	100
Malaysia	99	97	100	99	97	100
Norway (9)	96	92	99	96	91	99
Portugal	98	96	100	98	96	100
Qatar	97	96	99	97	96	99
Russian Federation	99	96	100	99	96	100
Singapore	98	96	100	98	96	100
Sweden	98	95	99	98	95	99
Turkey	99	97	100	99	97	100
United Arab Emirates	98	93	100	98	93	100
United States	98	94	100	98	94	100
International Average	98	95	100	98	95	100
Benchmarking Participants						
Ontario, Canada	95	93	98	95	93	98
Quebec, Canada	96	91	100	96	90	100
Moscow City, Russian Fed.	97	94	99	97	93	99
Abu Dhabi, UAE	99	98	100	99	98	100
Dubai, UAE	96	82	100	96	82	100

Within-Country Scoring Reliability for Human Scored Items in the TIMSS 2019 Problem Solving and Inquiry Tasks—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
Chile	94	91	99	94	90	99
Chinese Taipei	93	85	98	92	85	98
England	93	90	98	93	90	98
Finland	95	90	99	95	89	99
France	95	86	99	94	86	99
Georgia	94	88	99	93	88	99
Hong Kong SAR	100	100	100	100	100	100
Hungary	87	78	94	86	78	94
Israel	90	84	96	89	77	96
Italy	95	92	99	95	91	99
Korea, Rep. of	95	88	99	94	88	99
Lithuania	90	83	99	90	83	99
Malaysia	93	88	97	93	88	97
Norway (9)	92	87	99	91	84	99
Portugal	92	88	97	92	88	97
Qatar	94	92	98	94	91	98
Russian Federation	94	88	100	93	88	96
Singapore	95	88	100	94	88	100
Sweden	89	82	96	89	82	96
Turkey	95	88	100	95	86	100
United Arab Emirates	93	90	96	92	90	96
United States	93	87	100	93	87	99
International Average	93	88	98	93	87	98
Benchmarking Participants						
Ontario, Canada	91	80	98	90	80	98
Quebec, Canada	88	80	98	87	80	98
Moscow City, Russian Fed.	98	93	100	97	93	100
Abu Dhabi, UAE	94	89	99	94	89	99
Dubai, UAE	92	85	98	91	85	98