Examining Measurement Invariance Before and After the 2020 Pandemic

For Symposium Session: Impact of the Pandemic from Multiple Analytic Perspectives

Huan Wang

Dong-In Kim

Data Recognition Corporation

The 2022 Annual Meeting of the National Council on Measurement in Education

Introduction

A fundamental premise in assessment is that the underlying construct is equivalent across different groups of students and that this structure does not vary over years. The pandemic has potentially impacted opportunity to learn (Marion, 2020) and disrupted the internal structure of assessments in various ways.

This study evaluated the impact of the 2020 pandemic on the measurement invariance of a high-stakes assessment in mathematics. Past research (Kuhfeld et al., 2020) has suggested that students tended to have greater "COVID slide" in mathematics than in reading. Multigroup confirmatory factor analysis was conducted to explore whether measurement invariance held across the years prior to and immediately after the 2020 pandemic year: 2019 and 2021. In addition, the stability of measurement invariance was evaluated across different ethnic groups as well as groups of students from different district locales such as city and rural areas. Across groups of interest, configural invariance was confirmed as the first step, followed by examination of measurement invariance, including metric, scalar, and residual invariance. Results from the study contribute to a growing body of literature about the degree of impact from the pandemic on test score interpretations.

Method

Data

The study used operational test data from a large-scale assessment program in mathematics that tested students at grades 3 through 8. The assessments were administered under standardized conditions. Students were tested in five domains that represented grade-level content standards,

with six to eleven items in each domain. There were three item types: multiple-choice, short-answer, and technology-enhanced items. All items were scored dichotomously.

The data in the present study came from two test administrations: Spring 2019 and Spring 2021. The same operational test forms were used in the two administrations.

As the 2021 administration had lower participation rate than in a normal year with slight demographic shifts, the 2021 calibration sample was used instead of the 2021 census data. The calibration sample was created to be comparable in demographic characteristics to the 2019 test population by using the propensity score matching (PSM) method (Rosenbaum & Rubin, 1983). The covariates in the PSM model included student gender, ethnicity, disability status, English language proficiency status, economic status, and district locale as indicated in the National Center for Education Statistics database (https://nces.ed.gov/ccd/districtsearch)

To illustrate comparability of the 2019 and 2021 data sets, Tables 1 and 2 provide demographic distributions at grade 3 (an elementary school grade) and grade 7 (a middle school grade), respectively. The tables show the observed N counts and group percentages in each data set (Spring 2019 total test population and Spring 2021 calibration sample), by key demographic variables including gender, ethnicity, disability status, English language proficiency status, economic status, and district locale. It can be seen that the percentage difference between the two data sets on each demographic category was no greater than 0.7%. The pattern held at other grades as well.

Table 1.

Student Demographic Distributions in the 2019 and 2021 Data Sets, Grade 3

Group	2019 Census		2021 Calibration Sample		Difference % (2021 –2019)
	N	%	N	%	
All Students	61091	100.0	30474	100.0	
Gender:					
Female	29974	49.1	14891	48.9	-0.2
Male	31117	50.9	15583	51.1	0.2
Ethnicity:					
American Indian	726	1.2	297	1.0	-0.2
Asian	2512	4.1	1304	4.3	0.2
African American	6565	10.7	3235	10.6	-0.1
Hispanic	8295	13.6	4108	13.5	-0.1
White	40204	65.8	20141	66.1	0.3
Two or More	2745	4.5	1389	4.6	0.1
Disability:					
No	53064	86.9	26558	87.1	0.3
Yes	8027	13.1	3916	12.9	-0.3
Limited English Proficiency:					
No	55479	90.8	27717	91.0	0.1
Yes	5612	9.2	2757	9.0	-0.1
Economically Disadvantaged					
No	33672	55.1	16870	55.4	0.2
Yes	27419	44.9	13604	44.6	-0.2
District Locale:					
Non-Public	2989	4.9	1479	4.9	0.0
City	18707	30.6	9280	30.5	-0.2
Suburban	16667	27.3	8330	27.3	0.1
Town	11901	19.5	5969	19.6	0.1
Rural	10827	17.7	5416	17.8	0.0

Table 2.

Student Demographic Distributions in the 2019 and 2021 Data Sets, Grade 7

Group	2019 Census		2021 Calibration Sample		Difference % (2021 –2019)
	N	%	N	%	
All Students	63878	100.0	31812	100.0	
Gender:					
Female	31092	48.7	15287	48.1	-0.6
Male	32786	51.3	16525	51.9	0.6
Ethnicity:					
American Indian	805	1.3	357	1.1	-0.1
Asian	2493	3.9	1313	4.1	0.2
African American	6573	10.3	3153	9.9	-0.4
Hispanic	8672	13.6	4349	13.7	0.1
White	42845	67.1	21399	67.3	0.2
Two or More	2444	3.8	1241	3.9	0.1
Disability:					
No	56166	87.9	28050	88.2	0.2
Yes	7712	12.1	3762	11.8	-0.2
Limited English Proficiency:					
No	60272	94.4	29990	94.3	-0.1
Yes	3606	5.6	1822	5.7	0.1
Economically Disadvantaged					
No	36985	57.9	18438	58.0	0.1
Yes	26893	42.1	13374	42.0	-0.1
District Locale:					
Non-Public	2993	4.7	1484	4.7	0.0
City	18131	28.4	8946	28.1	-0.3
Suburban	18146	28.4	9121	28.7	0.3
Town	12796	20.0	6385	20.1	0.0
Rural	11812	18.5	5876	18.5	0.0

For the following analyses, student ethnic groups were coded as 1 (Hispanic), 2 (African American), 3 (White), 4 (Asian), and 5 (Others), and district locales were coded as 1 (City), 2 (Suburban), 3 (Town), 4 (Rural), and 5 (Others).

Table 3 through Table 8 present mean and standard deviation (STD) of observed domain raw scores in each data set (2019 and 2021) per grade by year, by ethnic groups within a year, and by district locales within a year.

Table 3.

Grade 3: Domain Raw Scores by Year, by Ethnic Groups Within Year, and by District Locales Within Year

Year	Group	Domain 1		Domain 2		Domain 3		Domain 4		Domain 5	
		Mean	***STD***	***Mean***	***STD***	***Mean***	***STD***	***Mean***	***STD***	***Mean***	***STD***
2019	Total	4.20	2.02	5.38	2.58	4.94	2.35	4.23	2.09	4.82	2.34
2021	Total	3.82	2.07	4.91	2.65	4.59	2.40	3.93	2.12	4.48	2.38
2019	Hispanic	3.42	1.94	4.47	2.44	4.05	2.33	3.52	1.89	4.05	2.20
2019	African American	2.68	1.85	3.41	2.31	3.04	2.22	2.79	1.76	3.32	2.10
2019	White	4.64	1.88	5.93	2.42	5.47	2.13	4.65	2.03	5.25	2.26
2019	Asian	4.14	2.07	5.35	2.61	4.95	2.39	4.18	2.08	4.89	2.41
2019	Other Ethnicity	3.83	2.03	4.93	2.56	4.51	2.36	3.84	2.07	4.44	2.30
2021	Hispanic	2.76	1.86	3.61	2.38	3.38	2.24	2.94	1.84	3.38	2.13
2021	African American	2.14	1.65	2.60	2.04	2.48	1.98	2.30	1.61	2.65	1.89
2021	White	4.36	1.93	5.61	2.47	5.22	2.18	4.45	2.03	5.04	2.27
2021	Asian	3.46	2.07	4.53	2.64	4.26	2.50	3.55	2.09	4.32	2.44
2021	Other Ethnicity	3.44	2.03	4.39	2.60	4.13	2.36	3.49	2.12	4.04	2.33
2019	City	3.61	2.07	4.67	2.63	4.26	2.45	3.69	2.08	4.27	2.35
2019	Suburban	4.76	1.89	6.03	2.48	5.51	2.14	4.72	2.09	5.39	2.29
2019	Town	4.46	1.95	5.74	2.46	5.22	2.24	4.50	2.04	5.02	2.30
2019	Rural	4.28	1.91	5.50	2.42	5.16	2.22	4.32	1.95	4.88	2.24
2019	Other Locale	3.41	1.96	4.36	2.41	4.04	2.38	3.54	1.91	4.08	2.17
2021	City	3.13	2.06	3.91	2.57	3.72	2.40	3.22	2.07	3.70	2.32
2021	Suburban	4.41	1.98	5.65	2.57	5.21	2.21	4.43	2.10	5.09	2.33
2021	Town	4.15	1.97	5.38	2.51	4.96	2.29	4.27	2.03	4.82	2.31
2021	Rural	4.00	1.95	5.32	2.50	4.97	2.27	4.21	2.00	4.74	2.29
2021	Other Locale	2.87	1.97	3.63	2.41	3.55	2.38	3.07	1.91	3.55	2.17

Table 4.

Grade 4: Domain Raw Scores by Year, by Ethnic Groups Within Year, and by District Locales Within Year

Year	Group	Domain 1		Domain 2		Domain 3		Domain 4		Domain 5	
		Mean	***STD***	***Mean***	***STD***	***Mean***	***STD***	***Mean***	***STD***	***Mean***	***STD***
2019	Total	3.51	1.90	4.88	2.39	4.53	2.47	4.59	2.85	4.66	2.06
2021	Total	3.21	1.92	4.65	2.32	4.19	2.49	4.31	2.86	4.42	2.02
2019	Hispanic	2.85	1.72	4.02	2.11	3.58	2.27	3.47	2.55	3.94	1.85
2019	African American	2.28	1.54	3.14	1.82	2.74	2.03	2.48	2.14	3.35	1.70
2019	White	3.88	1.86	5.39	2.34	5.05	2.37	5.22	2.76	5.07	2.02
2019	Asian	3.60	1.99	5.02	2.50	4.80	2.48	4.80	2.93	4.67	2.17
2019	Other Ethnicity	3.20	1.90	4.38	2.24	4.00	2.39	3.94	2.71	4.31	1.97
2021	Hispanic	2.51	1.68	3.76	1.96	3.08	2.17	3.00	2.46	3.64	1.81
2021	African American	1.84	1.36	2.93	1.65	2.20	1.76	2.01	1.85	3.04	1.60
2021	White	3.61	1.89	5.15	2.28	4.78	2.40	5.01	2.78	4.83	1.96
2021	Asian	3.03	1.95	4.49	2.42	3.99	2.50	3.99	2.87	4.24	2.13
2021	Other Ethnicity	2.87	1.86	4.10	2.16	3.63	2.35	3.59	2.73	4.03	1.94
2019	City	3.07	1.89	4.26	2.32	3.89	2.47	3.77	2.76	4.15	2.02
2019	Suburban	4.02	1.90	5.54	2.40	5.12	2.44	5.45	2.84	5.09	2.05
2019	Town	3.71	1.88	5.14	2.36	4.73	2.38	4.91	2.78	4.87	2.03
2019	Rural	3.47	1.79	4.92	2.25	4.72	2.34	4.70	2.65	4.83	2.00
2019	Other Locale	2.79	1.69	3.86	2.05	3.56	2.21	3.23	2.45	4.02	1.87
2021	City	2.59	1.80	3.95	2.16	3.38	2.38	3.31	2.64	3.86	1.95
2021	Suburban	3.70	1.92	5.21	2.35	4.76	2.47	5.04	2.87	4.80	2.03
2021	Town	3.53	1.90	4.98	2.32	4.61	2.42	4.87	2.86	4.70	2.00
2021	Rural	3.33	1.85	4.80	2.23	4.47	2.41	4.57	2.73	4.62	1.93
2021	Other Locale	2.52	1.68	3.72	1.99	3.08	2.12	2.88	2.37	3.72	1.83

Table 5.

Grade 5: Domain Raw Scores by Year, by Ethnic Groups Within Year, and by District Locales Within Year

Year	Group	Domain 1		Domain 2		Domain 3		Domain 4		Domain 5	
		Mean	***STD***	***Mean***	***STD***	***Mean***	***STD***	***Mean***	***STD***	***Mean***	***STD***
2019	Total	3.96	2.39	4.23	2.48	4.35	2.44	3.87	2.29	4.21	2.34
2021	Total	3.56	2.37	4.06	2.47	3.88	2.46	3.62	2.27	3.73	2.31
2019	Hispanic	3.06	2.14	3.44	2.18	3.44	2.27	3.00	1.96	3.41	2.17
2019	African American	2.26	1.84	2.64	1.85	2.61	2.02	2.37	1.63	2.55	1.85
2019	White	4.45	2.33	4.67	2.47	4.84	2.35	4.31	2.30	4.65	2.28
2019	Asian	4.08	2.56	4.47	2.58	4.85	2.48	4.15	2.36	4.67	2.39
2019	Other Ethnicity	3.61	2.33	3.80	2.40	3.86	2.39	3.46	2.18	3.85	2.28
2021	Hispanic	2.55	1.96	3.05	2.07	2.79	2.12	2.66	1.86	2.87	2.04
2021	African American	1.85	1.62	2.40	1.80	2.01	1.73	2.14	1.52	2.05	1.66
2021	White	4.10	2.36	4.58	2.46	4.44	2.41	4.10	2.27	4.21	2.27
2021	Asian	3.30	2.41	3.87	2.49	3.84	2.47	3.55	2.28	3.78	2.37
2021	Other Ethnicity	3.02	2.15	3.51	2.33	3.31	2.27	3.12	2.16	3.23	2.16
2019	City	3.36	2.34	3.73	2.39	3.66	2.42	3.30	2.19	3.59	2.32
2019	Suburban	4.53	2.40	4.78	2.55	4.84	2.42	4.42	2.34	4.69	2.31
2019	Town	4.20	2.34	4.47	2.49	4.67	2.37	4.08	2.27	4.45	2.29
2019	Rural	4.13	2.29	4.22	2.35	4.64	2.32	3.99	2.22	4.46	2.27
2019	Other Locale	2.73	2.03	3.13	2.09	3.49	2.24	2.94	1.91	3.35	2.11
2021	City	2.80	2.17	3.30	2.32	2.94	2.25	2.96	2.11	2.97	2.14
2021	Suburban	4.10	2.42	4.72	2.55	4.45	2.46	4.18	2.31	4.14	2.29
2021	Town	3.98	2.38	4.36	2.43	4.33	2.44	3.85	2.27	4.20	2.32
2021	Rural	3.81	2.28	4.25	2.32	4.31	2.35	3.83	2.19	4.06	2.27
2021	Other Locale	2.49	2.05	2.90	2.04	2.88	2.17	2.76	1.94	2.94	2.05

Table 6.

Grade 6: Domain Raw Scores by Year, by Ethnic Groups Within Year, and by District Locales Within Year

Year	Group	Domain 1		Domain 2		Domain 3		Domain 4		Domain 5	
		Mean	***STD***	***Mean***	***STD***	***Mean***	***STD***	***Mean***	***STD***	***Mean***	***STD***
2019	Total	5.15	2.90	2.66	1.86	5.24	2.72	3.17	1.86	5.27	2.33
2021	Total	4.66	2.81	2.40	1.75	4.72	2.72	2.86	1.78	4.89	2.21
2019	Hispanic	4.13	2.54	2.02	1.53	4.21	2.54	2.56	1.69	4.43	2.05
2019	African American	3.34	2.21	1.59	1.23	3.19	2.23	1.87	1.50	3.75	1.92
2019	White	5.67	2.89	2.97	1.91	5.81	2.61	3.51	1.81	5.71	2.29
2019	Asian	5.63	3.02	2.80	1.96	5.41	2.73	3.47	1.97	5.28	2.35
2019	Other Ethnicity	4.55	2.79	2.30	1.74	4.64	2.65	2.78	1.78	4.84	2.30
2021	Hispanic	3.43	2.23	1.76	1.35	3.45	2.32	2.13	1.50	3.99	1.86
2021	African American	2.96	1.95	1.49	1.14	2.73	1.96	1.59	1.31	3.50	1.76
2021	White	5.17	2.84	2.68	1.81	5.31	2.66	3.20	1.75	5.31	2.20
2021	Asian	5.21	3.05	2.49	1.89	4.90	2.75	3.07	1.95	4.92	2.27
2021	Other Ethnicity	3.98	2.61	2.06	1.57	4.04	2.57	2.51	1.68	4.42	2.05
2019	City	4.47	2.83	2.25	1.72	4.48	2.74	2.72	1.86	4.69	2.27
2019	Suburban	5.78	2.96	3.07	1.95	5.82	2.68	3.58	1.84	5.73	2.33
2019	Town	5.35	2.87	2.82	1.87	5.52	2.65	3.29	1.82	5.58	2.34
2019	Rural	5.25	2.79	2.67	1.81	5.47	2.57	3.23	1.76	5.38	2.23
2019	Other Locale	4.36	2.52	1.91	1.50	4.31	2.53	2.60	1.73	4.34	2.05
2021	City	3.86	2.59	2.03	1.58	3.88	2.60	2.38	1.68	4.33	2.12
2021	Suburban	5.28	2.94	2.74	1.88	5.26	2.79	3.24	1.82	5.32	2.27
2021	Town	4.75	2.76	2.47	1.73	4.93	2.64	2.94	1.75	5.00	2.16
2021	Rural	5.00	2.76	2.50	1.74	5.15	2.60	3.05	1.73	5.12	2.14
2021	Other Locale	3.92	2.43	1.86	1.45	3.87	2.48	2.27	1.65	4.20	1.96

Table 7.

Grade 7: Domain Raw Scores by Year, by Ethnic Groups Within Year, and by District Locales Within Year

Year	Group	Domain 1		Domain 2		Domain 3		Domain 4		Domain 5	
		Mean	***STD***	***Mean***	***STD***	***Mean***	***STD***	***Mean***	***STD***	***Mean***	***STD***
2019	Total	4.12	2.41	3.74	1.95	2.80	2.01	4.16	2.26	5.16	2.72
2021	Total	3.88	2.32	3.61	1.87	2.50	1.92	3.89	2.24	4.88	2.65
2019	Hispanic	3.30	2.02	3.19	1.65	1.98	1.72	3.36	2.13	4.18	2.39
2019	African American	2.63	1.70	2.65	1.48	1.37	1.41	2.34	1.83	3.16	2.06
2019	White	4.53	2.44	4.03	1.98	3.21	1.99	4.63	2.14	5.70	2.66
2019	Asian	4.40	2.61	4.02	2.15	2.97	2.14	4.31	2.28	5.28	2.86
2019	Other Ethnicity	3.60	2.30	3.42	1.83	2.33	1.89	3.64	2.26	4.54	2.63
2021	Hispanic	3.13	1.95	3.06	1.58	1.71	1.56	2.99	2.05	3.93	2.31
2021	African American	2.53	1.58	2.58	1.46	1.19	1.22	2.13	1.68	2.95	1.93
2021	White	4.25	2.36	3.88	1.89	2.87	1.92	4.36	2.16	5.39	2.61
2021	Asian	4.04	2.54	3.84	2.07	2.67	2.14	3.93	2.30	5.01	2.82
2021	Other Ethnicity	3.43	2.21	3.32	1.74	2.10	1.77	3.43	2.18	4.33	2.56
2019	City	3.57	2.29	3.35	1.84	2.27	1.94	3.51	2.28	4.44	2.67
2019	Suburban	4.69	2.52	4.15	2.07	3.34	2.05	4.71	2.20	5.77	2.76
2019	Town	4.22	2.40	3.84	1.91	2.94	1.95	4.39	2.17	5.42	2.63
2019	Rural	4.14	2.29	3.78	1.87	2.82	1.93	4.27	2.15	5.25	2.56
2019	Other Locale	3.43	2.12	3.15	1.74	1.96	1.72	3.33	2.12	4.20	2.45
2021	City	3.30	2.09	3.16	1.72	1.91	1.71	3.17	2.17	4.13	2.52
2021	Suburban	4.44	2.47	4.00	1.95	3.01	2.02	4.44	2.25	5.53	2.73
2021	Town	3.96	2.29	3.67	1.83	2.63	1.90	4.11	2.17	5.02	2.57
2021	Rural	3.93	2.25	3.72	1.83	2.60	1.84	4.10	2.13	5.05	2.53
2021	Other Locale	3.33	2.11	3.25	1.82	1.85	1.69	3.13	2.13	4.15	2.49

Table 8.

Grade 8: Domain Raw Scores by Year, by Ethnic Groups Within Year, and by District Locales Within Year

Year	Group	Domain 1		Domain 2		Domain 3		Domain 4		Domain 5	
		Mean	***STD***	***Mean***	***STD***	***Mean***	***STD***	***Mean***	***STD***	***Mean***	***STD***
2019	Total	4.25	2.54	4.51	2.50	4.45	2.59	3.31	2.26	3.47	1.86
2021	Total	3.86	2.42	4.30	2.44	4.10	2.48	2.84	2.17	3.38	1.84
2019	Hispanic	3.35	2.15	3.58	2.17	3.47	2.21	2.54	2.02	2.87	1.67
2019	African American	2.70	1.90	2.69	1.87	2.69	1.93	2.06	1.78	2.24	1.44
2019	White	4.64	2.55	4.95	2.47	4.90	2.57	3.63	2.25	3.78	1.84
2019	Asian	5.06	2.77	4.98	2.64	4.95	2.74	3.78	2.44	3.70	1.93
2019	Other Ethnicity	3.71	2.40	4.06	2.38	3.92	2.48	2.87	2.14	3.11	1.81
2021	Hispanic	3.03	2.03	3.37	2.04	3.19	2.08	2.16	1.84	2.77	1.59
2021	African American	2.39	1.65	2.59	1.72	2.46	1.69	1.70	1.53	2.12	1.41
2021	White	4.23	2.44	4.74	2.44	4.53	2.50	3.14	2.21	3.69	1.82
2021	Asian	4.47	2.73	4.62	2.53	4.40	2.67	3.26	2.38	3.45	1.90
2021	Other Ethnicity	3.35	2.27	3.70	2.34	3.48	2.30	2.38	2.05	2.99	1.84
2019	City	3.63	2.43	3.83	2.45	3.78	2.48	2.74	2.17	3.03	1.83
2019	Suburban	4.88	2.63	5.11	2.56	4.96	2.66	3.71	2.28	3.84	1.90
2019	Town	4.44	2.52	4.73	2.44	4.77	2.59	3.47	2.23	3.63	1.81
2019	Rural	4.18	2.40	4.52	2.33	4.51	2.41	3.42	2.22	3.53	1.75
2019	Other Locale	3.53	2.29	3.71	2.24	3.62	2.40	2.99	2.19	2.92	1.76
2021	City	3.22	2.21	3.65	2.33	3.41	2.28	2.32	1.98	2.90	1.78
2021	Suburban	4.38	2.56	4.89	2.54	4.55	2.60	3.16	2.28	3.79	1.89
2021	Town	4.07	2.41	4.45	2.39	4.39	2.47	3.01	2.17	3.48	1.81
2021	Rural	3.91	2.31	4.35	2.30	4.24	2.37	2.98	2.14	3.46	1.74
2021	Other Locale	3.29	2.23	3.51	2.20	3.49	2.42	2.62	2.13	2.86	1.74

Study Design

Multigroup confirmatory factor analysis was conducted to answer each of the following research questions:

1. Did measurement invariance of the test hold across years (2019 and 2021)?
2. Did measurement invariance hold across major ethnic groups in the 2019 and 2021 data sets, respectively?
3. Did measurement invariance hold across district locales in the 2019 and 2021 data sets?

The modeling was conducted using the lavaan R package (Rosseel, 2012). Based on the test design, a baseline model was specified at each grade. Figure 1 illustrates a general form of the baseline model. The actual set of domains may vary across grades. In the model, student total item scores on each content domain were used as observed values for the corresponding domain.

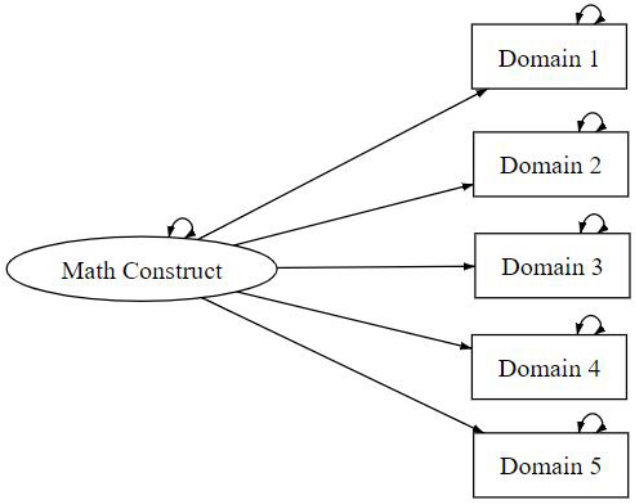

Figure 1. A General Form of the Baseline Model.

Before examining measurement invariance, the baseline model was separately tested for data fit in each group of interest to confirm configural invariance. Configural invariance means that the same model design holds across groups. Furthermore, configural invariance was tested by holding the same baseline model across groups in one single estimation. After configural invariance was confirmed, the baseline model was applied, and the *forward* (*sequential constraint imposition*) approach was conducted to test measurement invariance including metric,

scalar, and residual invariance. To test metric invariance, equal factor loadings were held across groups. Scalar invariance was tested by adding the constraints of equal intercepts across groups. Lastly, residual invariance was tested by adding the constraints of equal item uniquenesses (error variances) across groups.

The model fit (Brown, 2015) was evaluated using the chi-square test of fit ($\chi 2$) and fit indices including the comparative fit index (CFI), the root mean squared error of approximation (RMSEA), and the standardized root mean residual (SRMR). In addition, the model's ability to explain observed variance (R^2) was checked for reasonability. To compare models, the comparative fit index (CFI) change (ΔCFI) (Cheung & Rensvold, 2002) was examined.

Results

Table 9 through Table 11 present a summary of model fit statistics obtained from testing for measurement invariance a) between the 2019 and 2021 administrations at each grade, b) across ethnic groups in each test administration, and c) across district locales in each test administration. The model fit statistics on each separate student group (e.g., Hispanic at grade 3 in 2019) for evaluating configural invariance are also included.

The statistics in the tables indicate a reasonable model fit (CFI > 0.99, RMSEA < 0.08, and SRMR < 0.02) for all student groups of interest, except for Asian at grade 3 in 2021, where the CFI and RMSEA statistics are slightly worse (CFI = 0.989, RMSEA = 0.088, and SRMR = 0.015). It has been suggested that a CFI value greater than 0.95 indicates a good model-data fit (Hu and Bentler, 1999), a RMSEA value greater than 0.1 indicates poor fit (Fabrigar et al., 1999), and the acceptable range for the SRMR index is between 0 and 0.08 (Hu and Bentler, 1999). In addition, the R^2 statistics show that substantial variation of the domain scores in each group can be explained by model estimates (R^2 ranged from 0.4 to 0.8). Table 12 provides the R^2 statistics at grade 3 for an example. A similar pattern was observed in other grades as well. Furthermore, results from testing configural invariance show reasonable model fit in all cases (CFI > 0.99, RMSEA < 0.08, and SRMR < 0.02).

Overall, the model fit statistics support configural invariance of the baseline model over groups of interest (i.e., 2019 vs. 2021 administrations, ethnic groups within each administration, and district locales within each administration).

Table 9.

Testing for Measurement Invariance Between 2019 and 2021

Grade	Model Fit	Chi-square	df	CFI	ΔCFI	RMSEA	SRMR
3	*Year 2019*	1137.343	5	0.995		0.061	0.010
	Year 2021	582.081	5	0.995		0.062	0.010
	Configural Invariance	1719.425	10	0.995		0.061	0.010
	Metric Invariance	1736.389	14	0.995	0.000	0.052	0.011
	Scalar Invariance	1843.277	18	0.994	-0.001	0.047	0.011
	Residual Invariance	1868.526	23	0.994	0.000	0.042	0.011
4	*Year 2019*	747.625	5	0.996		0.048	0.009
	Year 2021	376.115	5	0.996		0.049	0.009
	Configural Invariance	1123.740	10	0.996		0.048	0.009
	Metric Invariance	1286.734	14	0.996	0.000	0.044	0.014
	Scalar Invariance	1525.693	18	0.995	-0.001	0.042	0.016
	Residual Invariance	1557.491	23	0.995	0.000	0.037	0.016
5	*Year 2019*	997.003	5	0.995		0.055	0.010
	Year 2021	365.586	5	0.996		0.047	0.008
	Configural Invariance	1362.589	10	0.995		0.053	0.009
	Metric Invariance	1389.700	14	0.995	0.000	0.045	0.010
	Scalar Invariance	2164.197	18	0.993	-0.002	0.050	0.016
	Residual Invariance	2251.130	23	0.992	-0.001	0.045	0.016
6	*Year 2019*	212.501	5	0.999		0.025	0.005
	Year 2021	101.711	5	0.999		0.024	0.005
	Configural Invariance	314.212	10	0.999		0.025	0.005
	Metric Invariance	528.424	14	0.998	-0.001	0.027	0.013
	Scalar Invariance	551.513	18	0.998	0.000	0.025	0.014
	Residual Invariance	619.534	23	0.998	0.000	0.023	0.016
7	*Year 2019*	524.586	5	0.997		0.040	0.007
	Year 2021	220.038	5	0.998		0.037	0.007
	Configural Invariance	744.624	10	0.997		0.039	0.007
	Metric Invariance	829.446	14	0.997	0.000	0.035	0.011
	Scalar Invariance	973.572	18	0.997	0.000	0.033	0.012
	Residual Invariance	1025.948	23	0.996	-0.001	0.030	0.013
8	*Year 2019*	1222.400	5	0.993		0.062	0.011
	Year 2021	673.724	5	0.992		0.065	0.012
	Configural Invariance	1896.124	10	0.993		0.063	0.012
	Metric Invariance	1942.560	14	0.993	0.000	0.054	0.013
	Scalar Invariance	2740.282	18	0.990	-0.003	0.057	0.018
	Residual Invariance	2803.680	23	0.990	0.000	0.051	0.019

Table 10.

Testing for Measurement Invariance Across Ethnic Groups, 2019 and 2021

Grade	Year	Model Fit	Chi-square	df	CFI	ΔCFI	RMSEA	SRMR
3	2019	*Hispanic*	149.023	5	0.994		0.059	0.012
		African American	92.259	5	0.995		0.052	0.011
		White	848.571	5	0.993		0.065	0.012
		Asian	25.162	5	0.998		0.040	0.007
		Others	84.139	5	0.993		0.068	0.013
		Configural Invariance	1199.154	25	0.994		0.062	0.012
		Metric Invariance	1797.476	41	0.991	-0.003	0.059	0.029
		Scalar Invariance	2478.103	57	0.987	-0.004	0.059	0.034
		Residual Invariance	2703.425	77	0.986	-0.001	0.053	0.037
3	2021	*Hispanic*	65.882	5	0.995		0.055	0.011
		African American	39.688	5	0.995		0.047	0.012
		White	414.039	5	0.994		0.064	0.011
		Asian	55.66	5	0.989		0.088	0.015
		Others	36.026	5	0.995		0.061	0.011
		Configural Invariance	611.295	25	0.994		0.062	0.012
		Metric Invariance	798.426	41	0.992	-0.002	0.055	0.024
		Scalar Invariance	1033.001	57	0.990	-0.002	0.053	0.028
		Residual Invariance	1174.951	77	0.988	-0.002	0.048	0.031
4	2019	*Hispanic*	104.553	5	0.995		0.048	0.011
		African American	30.740	5	0.998		0.027	0.008
		White	607.208	5	0.995		0.054	0.011
		Asian	32.686	5	0.997		0.045	0.008
		Others	29.468	5	0.998		0.037	0.008
		Configural Invariance	804.655	25	0.995		0.050	0.010
		Metric Invariance	1003.035	41	0.994	-0.001	0.043	0.017
		Scalar Invariance	1124.076	57	0.994	0.000	0.038	0.018
		Residual Invariance	1470.053	77	0.992	-0.002	0.038	0.022
4	2021	*Hispanic*	42.933	5	0.996		0.042	0.010
		African American	6.703	5	0.999		0.010	0.006
		White	331.232	5	0.994		0.056	0.012
		Asian	17.633	5	0.997		0.043	0.009
		Others	20.746	5	0.997		0.043	0.010
		Configural Invariance	419.247	25	0.995		0.050	0.011
		Metric Invariance	533.811	41	0.994	-0.001	0.044	0.019
		Scalar Invariance	641.142	57	0.993	-0.001	0.040	0.022
		Residual Invariance	1134.089	77	0.987	-0.006	0.047	0.028

Grade	Year	Model Fit	Chi-square	df	CFI	ΔCFI	RMSEA	SRMR
5	2019	*Hispanic*	124.801	5	0.994		0.052	0.011
		African American	57.803	5	0.996		0.039	0.011
		White	672.732	5	0.994		0.056	0.011
		Asian	55.071	5	0.994		0.062	0.011
		Others	59.967	5	0.995		0.056	0.010
		Configural Invariance	970.373	25	0.994		0.054	0.011
		Metric Invariance	1700.387	41	0.990	-0.004	0.056	0.028
		Scalar Invariance	1991.077	57	0.989	-0.001	0.051	0.031
		Residual Invariance	2737.821	77	0.984	-0.005	0.052	0.035
5	2021	*Hispanic*	28.370	5	0.998		0.032	0.008
		African American	26.440	5	0.996		0.036	0.011
		White	267.101	5	0.996		0.050	0.009
		Asian	35.121	5	0.993		0.068	0.012
		Others	12.081	5	0.999		0.029	0.007
		Configural Invariance	369.112	25	0.996		0.046	0.009
		Metric Invariance	521.043	41	0.994	-0.002	0.043	0.019
		Scalar Invariance	667.830	57	0.993	-0.001	0.041	0.022
		Residual Invariance	1403.255	77	0.984	-0.009	0.052	0.029
6	2019	*Hispanic*	91.950	5	0.996		0.044	0.011
		African American	57.069	5	0.996		0.039	0.012
		White	61.128	5	0.999		0.016	0.003
		Asian	20.072	5	0.998		0.034	0.008
		Others	12.913	5	0.999		0.021	0.005
		Configural Invariance	243.132	25	0.999		0.026	0.005
		Metric Invariance	1852.614	41	0.990	-0.009	0.058	0.041
		Scalar Invariance	2463.373	57	0.986	-0.004	0.057	0.044
		Residual Invariance	2876.774	77	0.984	-0.002	0.053	0.051
6	2021	*Hispanic*	19.638	5	0.998		0.026	0.008
		African American	30.917	5	0.993		0.040	0.015
		White	41.725	5	0.999		0.018	0.004
		Asian	41.978	5	0.991		0.075	0.017
		Others	19.388	5	0.997		0.042	0.009
		Configural Invariance	153.646	25	0.998		0.028	0.006
		Metric Invariance	886.960	41	0.989	-0.009	0.057	0.039
		Scalar Invariance	1190.764	57	0.986	-0.003	0.056	0.044
		Residual Invariance	1451.300	77	0.983	-0.003	0.053	0.051

Grade	Year	Model Fit	Chi-square	df	CFI	ΔCFI	RMSEA	SRMR
7	2019	*Hispanic*	49.388	5	0.998		0.032	0.007
		African American	10.711	5	0.999		0.013	0.005
		White	352.862	5	0.997		0.040	0.007
		Asian	57.128	5	0.994		0.064	0.011
		Others	9.093	5	0.999		0.016	0.004
		Configural Invariance	479.182	25	0.997		0.038	0.007
		Metric Invariance	1401.941	41	0.992	-0.005	0.051	0.030
		Scalar Invariance	2092.576	57	0.988	-0.004	0.053	0.037
		Residual Invariance	2848.199	77	0.983	-0.005	0.053	0.046
7	2021	*Hispanic*	25.626	5	0.998		0.031	0.008
		African American	17.559	5	0.996		0.029	0.010
		White	143.119	5	0.998		0.036	0.007
		Asian	12.168	5	0.998		0.033	0.007
		Others	19.555	5	0.996		0.043	0.009
		Configural Invariance	218.028	25	0.997		0.035	0.007
		Metric Invariance	592.843	41	0.993	-0.004	0.046	0.026
		Scalar Invariance	1006.884	57	0.988	-0.005	0.051	0.035
		Residual Invariance	1419.849	77	0.983	-0.005	0.053	0.043
8	2019	*Hispanic*	96.536	5	0.995		0.048	0.011
		African American	45.819	5	0.996		0.036	0.010
		White	928.038	5	0.992		0.065	0.013
		Asian	67.860	5	0.993		0.072	0.012
		Others	54.532	5	0.993		0.057	0.012
		Configural Invariance	1192.785	25	0.993		0.061	0.012
		Metric Invariance	1273.420	41	0.992	-0.001	0.049	0.015
		Scalar Invariance	2012.605	57	0.988	-0.004	0.052	0.022
		Residual Invariance	2581.139	77	0.984	-0.004	0.051	0.025
8	2021	*Hispanic*	68.021	5	0.991		0.056	0.015
		African American	28.395	5	0.993		0.039	0.013
		White	540.685	5	0.991		0.071	0.014
		Asian	15.042	5	0.998		0.040	0.008
		Others	11.297	5	0.998		0.029	0.007
		Configural Invariance	663.440	25	0.992		0.064	0.013
		Metric Invariance	700.367	41	0.991	-0.001	0.051	0.016
		Scalar Invariance	1135.177	57	0.986	-0.005	0.055	0.025
		Residual Invariance	1564.524	77	0.980	-0.006	0.056	0.030

Table 11.

Testing for Measurement Invariance Across District Locales, 2019 and 2021

Grade	Year	Model Fit	Chi-square	df	CFI	ΔCFI	RMSEA	SRMR
3	2019	*City*	364.811	5	0.995		0.062	0.010
		Suburban	312.149	5	0.994		0.061	0.011
		Town	203.32	5	0.995		0.058	0.010
		Rural	235.101	5	0.993		0.065	0.013
		Others	55.676	5	0.994		0.058	0.012
		Configural Invariance	1171.057	25	0.994		0.061	0.011
		Metric Invariance	1615.964	41	0.992	-0.002	0.056	0.024
		Scalar Invariance	2076.310	57	0.990	-0.002	0.054	0.027
		Residual Invariance	2294.781	77	0.989	-0.001	0.049	0.030
3	2021	*City*	139.350	5	0.996		0.054	0.009
		Suburban	155.185	5	0.995		0.060	0.010
		Town	165.198	5	0.992		0.073	0.013
		Rural	124.303	5	0.993		0.066	0.012
		Others	37.712	5	0.993		0.067	0.013
		Configural Invariance	621.748	25	0.994		0.063	0.011
		Metric Invariance	758.691	41	0.993	-0.001	0.054	0.020
		Scalar Invariance	932.232	57	0.992	-0.001	0.050	0.023
		Residual Invariance	1070.907	77	0.990	-0.002	0.046	0.025
4	2019	*City*	262.109	5	0.996		0.052	0.010
		Suburban	235.007	5	0.996		0.051	0.010
		Town	151.759	5	0.996		0.049	0.010
		Rural	76.049	5	0.998		0.035	0.008
		Others	23.811	5	0.997		0.034	0.009
		Configural Invariance	748.735	25	0.996		0.048	0.009
		Metric Invariance	849.540	41	0.996	0.000	0.039	0.014
		Scalar Invariance	1422.387	57	0.993	-0.003	0.043	0.019
		Residual Invariance	1520.231	77	0.992	-0.001	0.038	0.020
4	2021	*City*	92.952	5	0.997		0.044	0.009
		Suburban	146.908	5	0.995		0.057	0.011
		Town	56.069	5	0.997		0.041	0.009
		Rural	100.243	5	0.994		0.058	0.013
		Others	8.745	5	0.999		0.021	0.008
		Configural Invariance	404.917	25	0.996		0.049	0.010
		Metric Invariance	469.326	41	0.995	-0.001	0.041	0.016
		Scalar Invariance	699.441	57	0.993	-0.002	0.042	0.021
		Residual Invariance	908.133	77	0.991	-0.002	0.041	0.023

Grade	Year	Model Fit	Chi-square	df	CFI	ΔCFI	RMSEA	SRMR
5	2019	*City*	310.117	5	0.995		0.056	0.010
		Suburban	313.617	5	0.994		0.059	0.011
		Town	164.755	5	0.996		0.050	0.009
		Rural	128.160	5	0.996		0.046	0.009
		Others	56.376	5	0.992		0.058	0.014
		Configural Invariance	973.026	25	0.995		0.054	0.010
		Metric Invariance	1270.567	41	0.993	-0.002	0.048	0.020
		Scalar Invariance	1981.166	57	0.990	-0.003	0.051	0.025
		Residual Invariance	2235.302	77	0.988	-0.002	0.047	0.027
5	2021	*City*	147.727	5	0.995		0.055	0.010
		Suburban	115.859	5	0.996		0.050	0.009
		Town	39.287	5	0.998		0.033	0.006
		Rural	58.485	5	0.996		0.043	0.009
		Others	10.789	5	0.998		0.027	0.008
		Configural Invariance	372.147	25	0.996		0.047	0.009
		Metric Invariance	490.466	41	0.995	-0.001	0.041	0.018
		Scalar Invariance	1007.720	57	0.990	-0.005	0.051	0.026
		Residual Invariance	1304.038	77	0.987	-0.003	0.050	0.028
6	2019	*City*	138.808	5	0.998		0.038	0.007
		Suburban	23.243	5	0.999		0.014	0.003
		Town	31.475	5	0.999		0.020	0.004
		Rural	39.445	5	0.999		0.024	0.005
		Others	25.172	5	0.997		0.036	0.010
		Configural Invariance	258.144	25	0.999		0.027	0.005
		Metric Invariance	939.066	41	0.995	-0.004	0.041	0.028
		Scalar Invariance	1541.094	57	0.992	-0.003	0.045	0.031
		Residual Invariance	1696.468	77	0.992	0.000	0.040	0.034
6	2021	*City*	66.027	5	0.997		0.037	0.009
		Suburban	46.194	5	0.999		0.030	0.006
		Town	9.329	5	0.999		0.011	0.003
		Rural	5.498	5	0.999		0.004	0.003
		Others	19.210	5	0.996		0.042	0.013
		Configural Invariance	146.258	25	0.999		0.027	0.006
		Metric Invariance	383.875	41	0.996	-0.003	0.036	0.023
		Scalar Invariance	555.615	57	0.994	-0.002	0.037	0.026
		Residual Invariance	687.286	77	0.993	-0.001	0.035	0.030

Grade	Year	Model Fit	Chi-square	df	CFI	ΔCFI	RMSEA	SRMR
7	2019	*City*	136.745	5	0.998		0.038	0.007
		Suburban	189.143	5	0.997		0.045	0.008
		Town	89.415	5	0.998		0.036	0.007
		Rural	76.779	5	0.998		0.035	0.007
		Others	14.053	5	0.999		0.025	0.006
		Configural Invariance	506.134	25	0.997		0.039	0.007
		Metric Invariance	958.227	41	0.995	-0.002	0.042	0.023
		Scalar Invariance	1372.274	57	0.993	-0.002	0.043	0.026
		Residual Invariance	1686.722	77	0.991	-0.002	0.040	0.031
7	2021	*City*	46.482	5	0.998		0.031	0.006
		Suburban	111.635	5	0.996		0.048	0.009
		Town	42.317	5	0.998		0.034	0.007
		Rural	16.390	5	0.999		0.020	0.005
		Others	17.640	5	0.996		0.041	0.010
		Configural Invariance	234.465	25	0.997		0.036	0.007
		Metric Invariance	481.134	41	0.995	-0.002	0.041	0.023
		Scalar Invariance	697.913	57	0.992	-0.003	0.042	0.027
		Residual Invariance	886.898	77	0.990	-0.002	0.041	0.031
8	2019	*City*	289.221	5	0.994		0.057	0.011
		Suburban	440.202	5	0.992		0.069	0.012
		Town	293.458	5	0.992		0.067	0.013
		Rural	222.292	5	0.992		0.061	0.013
		Others	20.708	5	0.998		0.034	0.008
		Configural Invariance	1265.881	25	0.993		0.063	0.012
		Metric Invariance	1423.808	41	0.992	-0.001	0.052	0.018
		Scalar Invariance	1877.286	57	0.989	-0.003	0.050	0.021
		Residual Invariance	2166.082	77	0.988	-0.001	0.046	0.024
8	2021	*City*	234.599	5	0.989		0.073	0.015
		Suburban	191.122	5	0.993		0.064	0.012
		Town	108.564	5	0.994		0.056	0.011
		Rural	132.299	5	0.991		0.065	0.014
		Others	31.618	5	0.992		0.063	0.014
		Configural Invariance	698.202	25	0.992		0.066	0.013
		Metric Invariance	821.218	41	0.991	-0.001	0.055	0.021
		Scalar Invariance	1112.749	57	0.987	-0.004	0.054	0.025
		Residual Invariance	1322.498	77	0.985	-0.002	0.051	0.028

Table 12.

Domain Score Variation Explained by Model (R^2), Grade 3

Model Fit and Group	R^2				
	Domain 1	**Domain 2**	**Domain 3**	**Domain 4**	**Domain 5**
2019_Total	0.654	0.733	0.727	0.642	0.710
2021_Total	0.677	0.762	0.739	0.658	0.725
2019_Configural Invariance_Total	0.654	0.733	0.727	0.642	0.710
2021_Configural Invariance_Total	0.677	0.762	0.739	0.658	0.725
2019_Metric Invariance_Total	0.655	0.735	0.726	0.640	0.708
2021_Metric Invariance_Total	0.675	0.757	0.741	0.662	0.727
2019_Scalar Invariance_Total	0.656	0.736	0.725	0.640	0.708
2021_Scalar Invariance_Total	0.675	0.758	0.740	0.661	0.727
2019_Residual Invariance_Total	0.657	0.739	0.726	0.642	0.709
2021_ Residual Invariance_Total	0.673	0.752	0.739	0.658	0.723
2019_Hispanic	0.588	0.703	0.706	0.546	0.654
2019_African American	0.578	0.705	0.687	0.529	0.640
2019_White	0.613	0.692	0.683	0.624	0.697
2019_Asian	0.674	0.736	0.719	0.641	0.741
2019_Other Ethnicity	0.628	0.722	0.725	0.616	0.686
2019_Configural Invariance_Hispanic	0.588	0.703	0.706	0.546	0.654
2019_Configural Invariance_African American	0.613	0.692	0.683	0.624	0.697
2019_Configural Invariance_White	0.578	0.705	0.687	0.529	0.640
2019_Configural Invariance_Asian	0.628	0.722	0.725	0.616	0.686
2019_Configural Invariance_Other Ethnicity	0.674	0.736	0.719	0.641	0.741
2019_Metric Invariance_Hispanic	0.586	0.694	0.654	0.597	0.674
2019_Metric Invariance_African American	0.616	0.696	0.699	0.601	0.689
2019_Metric Invariance_White	0.573	0.690	0.633	0.592	0.658
2019_Metric Invariance_Asian	0.620	0.720	0.704	0.623	0.708
2019_Metric Invariance_Other Ethnicity	0.658	0.740	0.717	0.656	0.740
2019_Scalar Invariance_Hispanic	0.594	0.695	0.663	0.597	0.660
2019_Scalar Invariance_African American	0.624	0.696	0.705	0.600	0.672
2019_Scalar Invariance_White	0.581	0.692	0.638	0.592	0.632
2019_Scalar Invariance_Asian	0.629	0.721	0.712	0.623	0.695
2019_Scalar Invariance_Other Ethnicity	0.667	0.741	0.724	0.655	0.725
2019_Residual Invariance_Hispanic	0.617	0.698	0.694	0.602	0.669
2019_Residual Invariance_African American	0.617	0.698	0.694	0.602	0.669
2019_Residual Invariance_White	0.586	0.670	0.666	0.570	0.639
2019_Residual Invariance_Asian	0.649	0.727	0.723	0.635	0.699
2019_Residual Invariance_Other Ethnicity	0.666	0.741	0.738	0.652	0.714

Model Fit and Group	R^2				
	Domain 1	Domain 2	Domain 3	Domain 4	Domain 5
2021_Hispanic	0.607	0.710	0.692	0.545	0.657
2021_African American	0.536	0.671	0.631	0.475	0.602
2021_White	0.628	0.713	0.690	0.627	0.691
2021_Asian	0.653	0.772	0.745	0.610	0.765
2021_Other Ethnicity	0.629	0.760	0.715	0.644	0.729
2021_Configural Invariance_Hispanic	0.628	0.713	0.690	0.627	0.691
2021_Configural Invariance_African American	0.607	0.710	0.692	0.545	0.657
2021_Configural Invariance_White	0.629	0.760	0.715	0.644	0.729
2021_Configural Invariance_Asian	0.653	0.772	0.745	0.610	0.765
2021_Configural Invariance_Other Ethnicity	0.536	0.671	0.631	0.475	0.602
2021_Metric Invariance_Hispanic	0.626	0.716	0.702	0.610	0.689
2021_Metric Invariance_African American	0.605	0.703	0.652	0.591	0.670
2021_Metric Invariance_White	0.639	0.757	0.710	0.638	0.734
2021_Metric Invariance_Asian	0.665	0.777	0.718	0.646	0.752
2021_Metric Invariance_Other Ethnicity	0.535	0.664	0.582	0.535	0.608
2021_Scalar Invariance_Hispanic	0.633	0.717	0.707	0.608	0.675
2021_Scalar Invariance_African American	0.611	0.705	0.659	0.590	0.657
2021_Scalar Invariance_White	0.646	0.759	0.716	0.636	0.723
2021_Scalar Invariance_Asian	0.670	0.779	0.724	0.643	0.735
2021_Scalar Invariance_Other Ethnicity	0.543	0.666	0.587	0.533	0.589
2021_Residual Invariance_Hispanic	0.635	0.725	0.701	0.614	0.681
2021_Residual Invariance_African American	0.607	0.701	0.676	0.586	0.655
2021_Residual Invariance_White	0.666	0.752	0.729	0.646	0.710
2021_Residual Invariance_Asian	0.683	0.766	0.744	0.663	0.726
2021_Residual Invariance_Other Ethnicity	0.516	0.618	0.589	0.494	0.567
2019_City	0.658	0.752	0.739	0.644	0.707
2019_Suburban	0.626	0.716	0.701	0.638	0.709
2019_Town	0.639	0.696	0.705	0.619	0.709
2019_Rural	0.598	0.677	0.687	0.598	0.673
2019_Other Locales	0.613	0.713	0.723	0.549	0.653
2019_Configural Invariance_City	0.598	0.677	0.687	0.598	0.673
2019_Configural Invariance_Suburban	0.639	0.696	0.705	0.619	0.709
2019_Configural Invariance_Town	0.626	0.716	0.701	0.638	0.709
2019_Configural Invariance_Rural	0.658	0.752	0.739	0.644	0.707
2019_Configural Invariance_Other Locale	0.613	0.713	0.723	0.549	0.653
2019_Metric Invariance_City	0.599	0.684	0.679	0.607	0.666
2019_Metric Invariance_Suburban	0.633	0.709	0.709	0.615	0.697
2019_Metric Invariance_Town	0.638	0.711	0.727	0.606	0.694
2019_Metric Invariance_Rural	0.652	0.746	0.722	0.658	0.723
2019_Metric Invariance_Other Locale	0.600	0.714	0.664	0.598	0.678

Model Fit and Group	R^2				
	Domain 1	Domain 2	Domain 3	Domain 4	Domain 5
2019_Scalar Invariance_City	0.605	0.684	0.678	0.607	0.659
2019_Scalar Invariance_Suburban	0.639	0.709	0.710	0.615	0.689
2019_Scalar Invariance_Town	0.642	0.712	0.727	0.605	0.688
2019_Scalar Invariance_Rural	0.658	0.747	0.723	0.658	0.717
2019_Scalar Invariance_Other Locale	0.605	0.714	0.665	0.598	0.670
2019_Residual Invariance_City	0.614	0.696	0.690	0.599	0.669
2019_Residual Invariance_Suburban	0.633	0.712	0.707	0.618	0.686
2019_Residual Invariance_Town	0.629	0.709	0.703	0.614	0.682
2019_Residual Invariance_Rural	0.667	0.742	0.737	0.653	0.717
2019_Residual Invariance_Other Locale	0.621	0.702	0.696	0.606	0.675
2021_City	0.688	0.765	0.735	0.645	0.726
2021_Suburban	0.657	0.740	0.713	0.652	0.716
2021_Town	0.618	0.727	0.722	0.639	0.701
2021_Rural	0.624	0.704	0.693	0.600	0.683
2021_Other Locales	0.637	0.750	0.715	0.570	0.657
2021_Configural Invariance_City	0.624	0.704	0.693	0.600	0.683
2021_Configural Invariance_Suburban	0.618	0.727	0.722	0.639	0.701
2021_Configural Invariance_Town	0.657	0.740	0.713	0.652	0.716
2021_Configural Invariance_Rural	0.688	0.765	0.735	0.645	0.726
2021_Configural Invariance_Other Locale	0.637	0.750	0.715	0.570	0.657
2021_Metric Invariance_City	0.630	0.705	0.688	0.608	0.677
2021_Metric Invariance_Suburban	0.632	0.731	0.714	0.636	0.695
2021_Metric Invariance_Town	0.660	0.734	0.735	0.631	0.709
2021_Metric Invariance_Rural	0.677	0.768	0.725	0.654	0.733
2021_Metric Invariance_Other Locale	0.626	0.750	0.667	0.609	0.681
2021_Scalar Invariance_City	0.629	0.709	0.688	0.607	0.670
2021_Scalar Invariance_Suburban	0.634	0.734	0.715	0.635	0.689
2021_Scalar Invariance_Town	0.660	0.738	0.735	0.630	0.703
2021_Scalar Invariance_Rural	0.678	0.771	0.725	0.654	0.728
2021_Scalar Invariance_Other Locale	0.626	0.753	0.667	0.609	0.673
2021_Residual Invariance_City	0.637	0.730	0.703	0.617	0.687
2021_Residual Invariance_Suburban	0.647	0.738	0.712	0.628	0.696
2021_Residual Invariance_Town	0.653	0.743	0.717	0.634	0.701
2021_Residual Invariance_Rural	0.669	0.756	0.731	0.649	0.716
2021_Residual Invariance_Other Locale	0.632	0.725	0.698	0.612	0.681

To evaluate measurement invariance, the CFI difference (ΔCFI) was examined by subtracting the CFI statistic of a more constrained model from that of the previous model which was less constrained. For example, e.g., when examining invariance between 2019 and 2021, the CFI at grade 6 for the baseline model from testing configural invariance (parameters were free to estimate) was 0.999, and the CFI for the baseline model from testing metric invariance (factor loadings were constrained to be the same across 2019 and 2021) was 0.998; therefore, the corresponding ΔCFI was 0.001 (0.998 − 0.999 = -0.001). The obtained ΔCFI value was well above the threshold of -0.01 (Cheung & Rensvold, 2002), which suggests metric invariance. As shown in Table 9 through Table 11, across grades and groups the observed ΔCFI values (see Table 9 through Table 11) were all above the threshold from testing for metric, scale, and residual invariance. The results support measurement invariance (metric, scalar, and residual) across years (2019 vs. 2021), across ethnic groups within each year, and across district locales within each year.

Discussion

The present study is intended to provide a big picture about the assessment, such as if the assessment construct still holds and if the reporting category design still has similar meaning across the groups under study. Results from the study provide supporting evidence that measurement invariance (metric, scalar, and residual invariance) held reasonably well across the years (2019 vs. 2021) for the assessment under study, despite potential impact of the 2020 pandemic on student learning. The analysis on metric invariance provides evidence for comparable domain relationship in the measurement across years, across ethnic groups, and across district locales; the analysis on scalar invariance provides evidence for comparable domain difficulty in the measurement across the groups of interest; and the residual invariance suggests comparable patterns of domain score variability in the measurement across the groups of interest.

The general picture, however, does not exclude possibilities that score interpretations may vary for groups of smaller sample sizes (e.g., Asian) and vary locally. Certain districts may have been hit harder than others; similarly, certain student groups such as African American students in an urban area may have had a different academic experience in 2020, which impact was not investigated in the present study.

It should also be noted that the baseline model showed slightly worse model fit for Asian at grade 3 in 2021. In addition, although above the threshold (ΔCFI = -0.01), the ΔCFI values observed in the following cases were close to the threshold: At grade 5 in 2021, the ΔCFI value was -0.009 on testing for residual invariance across ethnic groups; at grade 6 in both the 2019 and 2021 administrations, the ΔCFI value was also -0.009 on testing for metric invariance across ethnic groups. Those groups may deserve further investigation in the future.

Given that the 2021 calibration sample was used in the study, instead of the 2021 census data, it may also be interesting to examine if the 2021 calibration sample and the 2021 census data would lead to different analysis results.

Another limitation of the present study is that the analysis only focused on the content domain score relationship in the assessment. Further analysis is recommended to examine item level performance and evaluate if measurement invariance also held within each content domain.

References

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). The Guilford Press.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233–255.

Fabrigar L. R., MacCallum R. C., Wegener D. T., & Strahan E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272–299.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55.

Kuhfeld, M., Tarasawa, B., Johnson, A., Ruzek, E., and Lewis, K. (2020). *Learning during COVID-19: Initial findings on students' reading and math achievement and growth*. NWEA. Available online at: https://www.ewa.org/sites/main/files/file-attachments/learning_during_covid-19_brief_nwea_nov2020_final.pdf?1606835922

Marion, S. (October, 2020). Using Opportunity to Learn Data to Support Educational Equity. Center for Assessment. Retrieved from https://www.nciea.org/sites/default/files/inline-files/CFA-Marion.OTL_.Indicators_0.pdf

Rosenbaum, P.R., & Rubin, D.B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *17*(1), 41–55.

Rosseel Y (2012). "lavaan: An R Package for Structural Equation Modeling." *Journal of Statistical Software*, 48(2), 1–36. https://www.jstatsoft.org/v48/i02/