HumRRO
HUMAN RESOURCES RESEARCH ORGANIZATION

2022 No. 134

Progress Since COVID-19's Onset: Impact Study
Final Report

Prepared for: Kentucky Department of Education
Office of Assessment and Accountability
300 Sower Boulevard
Frankfort, KY 40601

Prepared under: Contract #1900004339

Authors: Gertrudes Velasquez
Arthur A. Thacker
Emily R. Dickinson

Date: January 25, 2023

Progress Since COVID-19's Onset: Impact Study

Table of Content

List of Tables

List of Figures

Progress Since COVID-19's Onset: Impact Study

Introduction

Kentucky typically administers its summative assessment in the spring semester to grades 3-8 and high school for math and reading, and once in elementary, middle, and high school for writing, science, and social studies. In spring 2019, Kentucky students were administered single K-PREP test forms for mathematics and reading. Students' ability estimates on these subjects were transformed to scale scores using scoring tables generated during a previous administration year. Scale scores ranged from 100 – 300 (Bynum & Thacker, 2020).

The 2019-2020 school year saw unprecedented long-term school closures across the Commonwealth of Kentucky, due to the COVID-19 pandemic. As a result of federal accountability waivers, Kentucky students did not take statewide operational assessments in 2020. Samples of students from different grades, however, took multiple test forms consisting of reading and mathematics field test (FT) items administered in early spring 2020. These items were administered to students in subsequent operational summative assessments.

Starting with the 2021 summative assessments, the Kentucky Department of Education (KDE) implemented several changes to its testing program to reduce students' test-taking time while maintaining sufficient rigor for accurate student-level reporting (Thacker, 2021). 2020 field test items were used operationally on multiple test forms for the spring 2022 Kentucky Summative Assessments (KSA) to provide student scores and performance ratings under the revised testing program.

The impacts of non-traditional instruction (NTI) on student performance implemented during the Covid-19 pandemic are not yet fully understood. In this study, we conduct a comparison of Kentucky student performance on the 2022 mathematics (*Math*) and reading (*Reading*) summative assessments with the corresponding 2019 assessments to evaluate student progress since COVID-19's onset.

Methods

KDE provided HumRRO with student-level data files for 2020 reading and mathematics field test items and for 2019 and 2022 operational assessments. We applied exclusion rules, as described in KDE's 2019 and 2022 psychometric analyses specifications documents (*Kentucky Spring 2019 Psychometric Analysis Specifications v1.3* and *KSA Spring 2022 Psychometric Analysis Specifications v1.3*), to select samples of student responses to include in calibration and equating analyses. We verified n-counts by subject and grade for each school year with results from psychometric analyses from these years.

We transformed 2019 students' ability estimates for reading and mathematics to a reporting scale with range from 100 – 300 using previously generated scale score (SS) tables. This was possible because 2019 reading and mathematics test forms were reused from a previous administration year.

We were not able to directly equate 2022 operational assessments to 2019 operational assessments because there were no common items on these assessments. However, the 2022 operational assessments did share items in common with the 2020 field tests. And, because many students who participated in the 2020 field tests also participated in the 2019 operational

assessments, this created a logical chain for putting the 2019 and 2022 assessments on comparable scales.

We first compared students who took both 2019 operational tests and 2020 field test items to the full population of students who took the 2019 operational tests. This step ensured that our analyses were not based on a sample of students who were substantially different from the full population of Kentucky students. To accomplish this, we matched students who took both the 2019 operational assessments and the 2020 field test items. We then compared 2019 mean ability estimates and standard deviations (SDs) of the matched samples with that of the entire population of students tested in 2019[1].

Next, we estimated parameters for 2020 field test items for all test forms in a grade concurrently, using item scores from the matched sample of students. This equated the 2020 field test forms for a given grade so that test scores across forms were interchangeable in terms of difficulty. We then used these 2020 field test item parameter estimates as input parameters in an anchored item equating process to place 2022 operational tests onto the same measurement scale as the 2020 field test items. We transformed 2022 students' ability estimates to the 2019 reporting scale using 2019 transformation constants specified in KDE's 2019 psychometric analyses specifications document (*Kentucky Spring 2019 Psychometric Analysis Specifications v1.3*). This yielded scale scores from 2019 and 2022 that were comparable at the state level.

We computed scale score summary statistics (means and SDs) and standardized mean differences (effect sizes, specifically Cohen's *d*) to examine overall differences in 2019 and 2022 student performance. We also examined mean scale score differences by students' gender (*Gender*), ethnicity (*Ethnicity*), individual education plan (*IEP*), English language proficiency (*EL*), and eligibility for free/reduced lunch (*Lunch*). We examined between-groups, within-year differences (e.g., scale score differences between female and male students on 2019 assessments) and within-group, between-years differences (e.g., scale score differences between female groups in 2019 and 2022). Our analyses for the *Ethnicity* variable were restricted to group performance for White, Black, and Hispanic students because of small sample sizes for other available groups. Finally, we summarized key observations on overall and group performances on KSA and National Assessment of Educational Progress (NAEP) assessments.

---

[1] We repeated this procedure for matched samples between 2020 field test forms and the 2019 operational assessments. We found no systematic differences between test forms and operational assessments.

Results

Comparison between Students' Scale Scores on 2019 Operational Tests and 2020 Field Tests

Table 1 presents mean scale scores and standard deviations (SDs), on the KSA 2019 100 – 300 reporting scale, of full populations of students who took the 2019 operational tests and subpopulations of students who took both 2019 operational tests and 2020 field test items. Mean scale score differences ranged from 0.40 to 1.14 points for mathematics and from 0.50 to 0.93 for reading, with average scale scores for the matched samples being consistently larger than the average scores for the full student population across subjects and grades. Figure 1 presents scale score distributions for matched samples and full populations by subject and grade. Vertical dashed lines on each plot represent mean scale score estimates for each of the groups. In general, matched sample ability distributions were representative of the full population of students who took the 2019 operational tests. Consequently, we did not adjust for sampling bias on the 2020 field test forms.

Table 1. Comparison between mean scale score estimates on math and reading 2019 operational tests and 2020 field tests

Subject	Grade (2019)	2019 OP Count	2019 OP Mean Scale Scores	2019 OP Standard Deviation	2020 FT Count	2020 FT Mean Scale Scores	2020 FT Standard Deviation	Scale Score differences
Math	3	48413	209.67	20.93	21924	210.53	20.47	0.86
Math	4	50655	209.47	19.28	11302	209.87	19.17	0.40
Math	5	51137	210.43	19.78	22727	211.20	19.47	0.77
Math	6	51132	209.72	18.61	22696	210.71	18.19	0.99
Math	7	49751	210.24	18.41	11028	211.38	18.14	1.14
Reading	3	48394	209.8	19.28	21404	210.44	18.75	0.64
Reading	4	50651	209.96	16.43	10910	210.46	15.71	0.50
Reading	5	51073	212.65	16.88	22535	213.41	16.02	0.76
Reading	6	51047	212.26	16.45	22506	213.19	15.56	0.93
Reading	7	49735	211.42	16.20	11038	212.21	15.45	0.79

Note. 2019 OP refers to all students who participated in the 2019 operational administration. 2020 FT refets to students who participated in the 2019 operational administration and also in the 2020 field test administration.
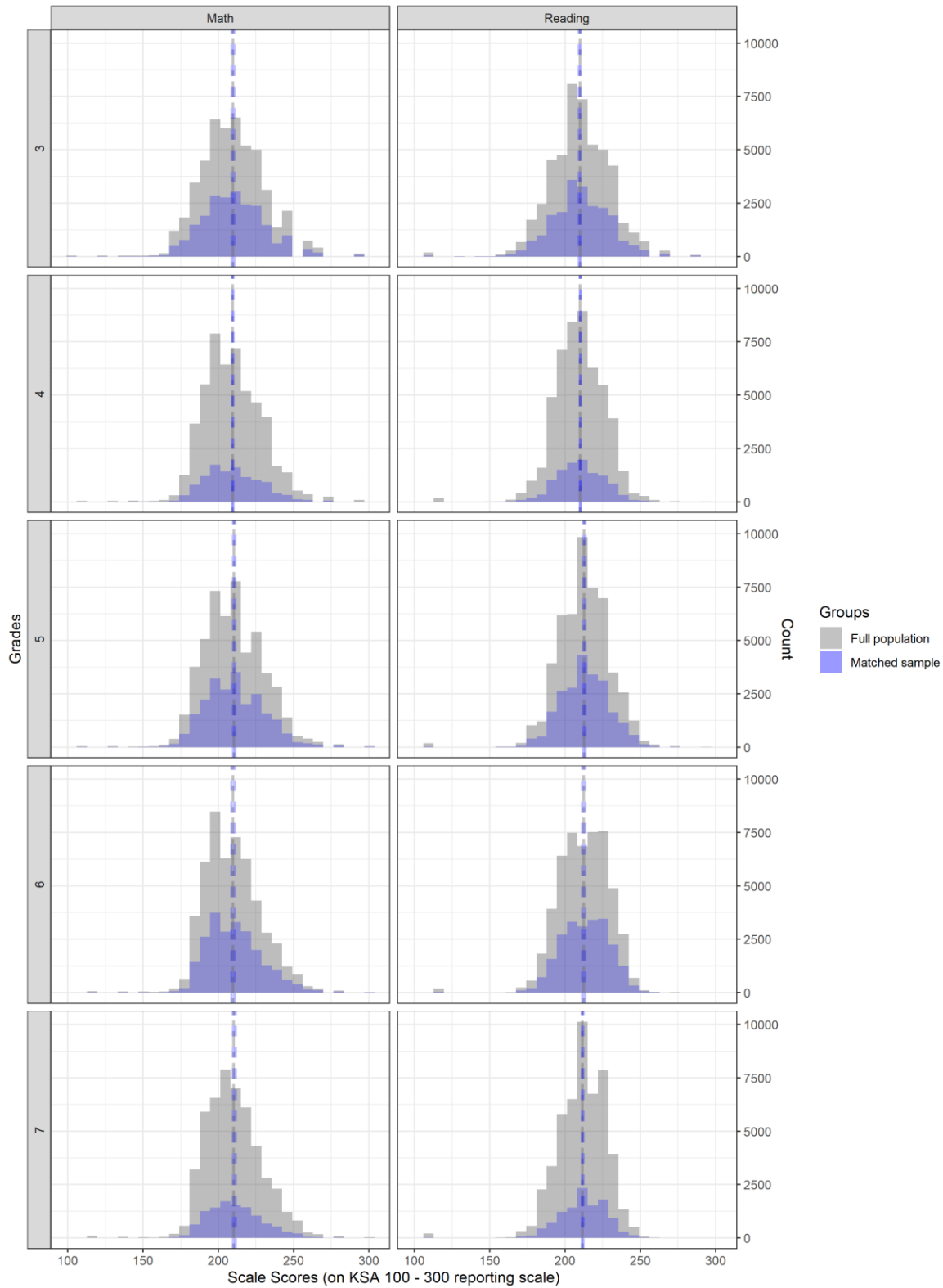
Figure 1. Scale score distributions for the full population of students who took the 2019 operational tests and subpopulations of students who took both the math and reading 2019 operational tests and 2020 field tests

Overall Mean Scale Score Comparisons

Table 2 and Figures 2 and 3 present scale scores summary statistics across grades for 2019 and 2022 mathematics and reading operational data. Across grades and subjects, mean scale scores on the 2022 operational tests for both subjects were consistently smaller than mean scale scores on the 2019 operational tests. Beginning at grade 4, 2022 – 2019 mean scale scores differences for math generally increased as grade level increased. These differences ranged from 9.69 points (in math grade 4) to 20.96 points (in math grade 6) on a 200-point scale, equivalent to a range in standardized mean differences (Cohen's d) from 0.5 to 1.20 standard deviations. For reading, mean scale differences decreased between grades 3 and 5 but increased for grades beyond grade 5. 2022 – 2019 reading mean scale scores differences ranged from 7.94 points (in reading grade 5) to 18.49 points (in reading grade 7) on a 200-point scale. These differences are equivalent to a range in standardized mean differences from 0.46 to 1.16 standard deviations.

Table 2. Scale scores summary statistics for 2019 and 2022 mathematics and reading operational data

Subject	Grade	2019 OP Count	2019 OP Mean Scale Scores	2019 OP Standard Deviation	2022 OP Count	2022 OP Mean Scale Scores	2022 OP Standard Deviation	Mean Scale Score differences (2022 – 2019)	Cohen's d
Mathematics	3	48617	209.55	21.06	46795	193.53	20.43	-16.02	-0.77
Mathematics	4	50731	209.58	19.22	43201	199.89	19.92	-9.69	-0.50
Mathematics	5	51575	211.26	19.93	46667	192.00	19.65	-19.26	-0.97
Mathematics	6	51635	209.75	18.42	46956	188.79	16.33	-20.96	-1.20
Mathematics	7	49971	210.22	18.07	49174	192.29	14.41	-17.93	-1.10
Mathematics	8	49278	208.84	18.29	50176	188.64	16.16	-20.20	-1.17
Reading	3	48450	210.20	18.26	46570	195.12	17.37	-15.08	-0.85
Reading	4	50568	210.45	15.47	42985	198.63	17.51	-11.82	-0.72
Reading	5	51399	212.23	15.75	46457	204.29	18.88	-7.94	-0.46
Reading	6	51486	212.66	15.40	46795	196.88	16.86	-15.78	-0.98
Reading	7	49837	212.11	15.44	49000	193.62	16.41	-18.49	-1.16
Reading	8	49127	214.03	15.76	50004	198.95	15.54	-15.08	-0.96

Note. OP=Operational. Differences in sample sizes reported in Table 1 and samples with scale scores reported in Table 2 are a result of assigning scale scores to students who reported total raw scores on operational assessments but were excluded from the calibration samples.
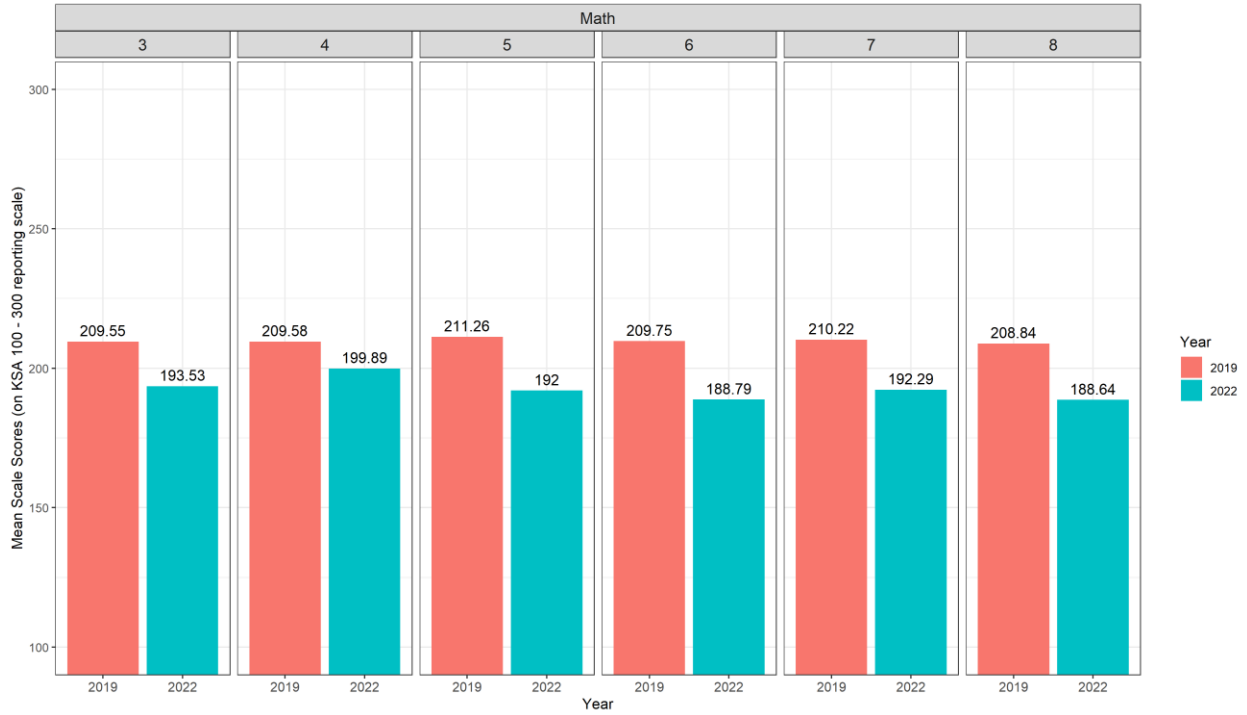
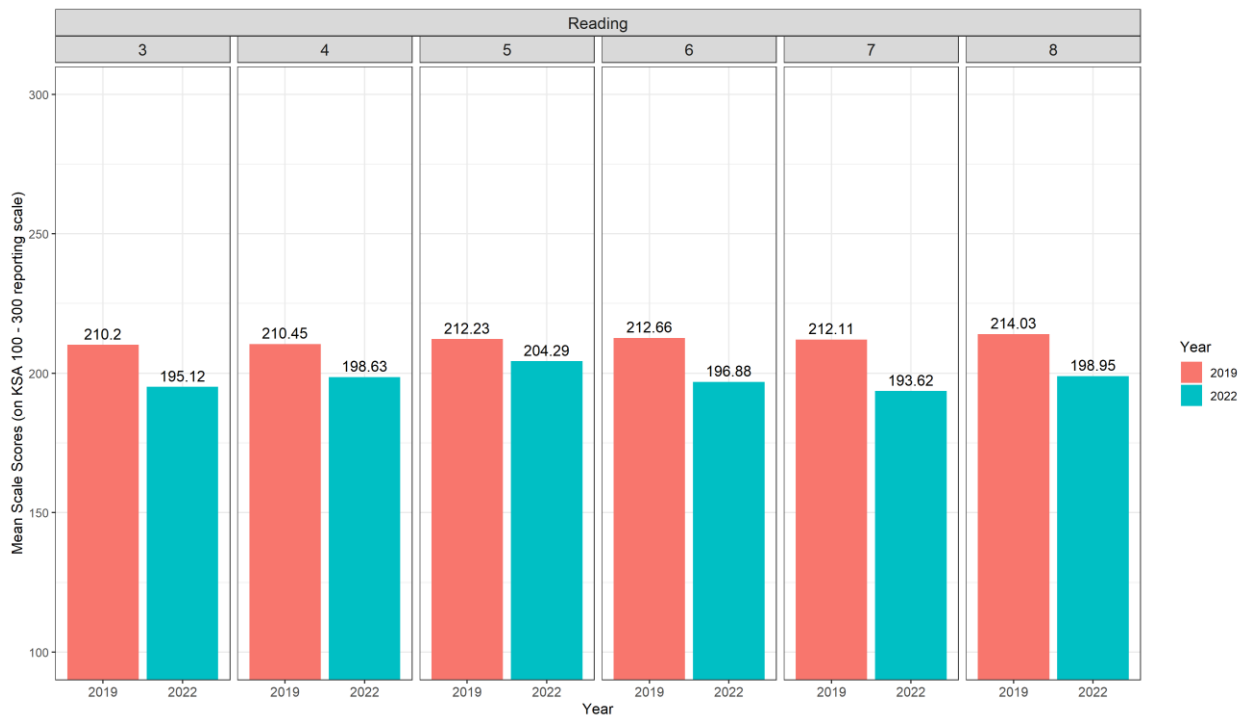**Figure 2. 2019 and 2022 mean scale scores for math grades**

Figure 3. 2019 and 2022 mean scale scores for reading grades

Mean Scale Score Comparisons by Group

We examined mathematics and reading mean scale score differences between groups identified by gender, ethnicity, English language proficiency, individual education plan, and eligibility for free/reduced lunch. For each variable, we compared group mean scale scores within grades and between years (e.g., 2019 and 2022 math grade 3 mean scale score differences) and across grades and within years (e.g., 2019 math grades 3 – 8 mean scale score differences, 2022 math grades 3 – 8 mean scale score differences, respectively). We also compared group performance across grades and years (e.g., 2019 math grades 3 – 8 mean scale score differences with 2022 math grades 3 – 8 mean scale score differences). Figures 4 through 8 depict these comparisons graphically. Tables 4a through 4e in Appendix A present the same information in tabular form.

Between-groups comparisons

Comparisons based on gender

We compared mathematics and reading mean scale scores for female and male students. The male group served as the reference group in these comparisons. To facilitate interpretation of group comparisons from the plots in the subsequent figures below, consider the graph for mathematics grade 6 in Figure 4 below. The horizontal dashed line near the center of the panel represents no difference between groups (that is, on the vertical axis, a scale score difference of zero). Shaded circles at the end of the line graph (solid line) indicate scale score differences between groups, one for 2019 and one for 2022. A shaded circle falling directly on the horizontal dashed line would indicate that male and female mean mathematics scale scores were the same in a particular year. Positive scale score differences (values above the horizontal dashed line) indicate that mean scale scores for the focus group (females in this comparison, for example) were larger than the mean scores for the reference group (males in this comparison), and vice versa. In 2019, for example, the mean math scale score for female students was nearly 2 points higher than the male mean math score. In 2022, however, the mean math score for females was slightly lower (about ½ point) than the male score (notice that the shaded circle corresponding to the 2022 scale score difference shifted below the horizontal dashed line). Connecting the shaded circles with a line provides a slope that describes the change in mean score differences, in this case the gap between males and females declined and even changed direction. Similar interpretations follow for other group comparisons.

2019 and 2022 mathematics mean scale scores for female students in grades 3 – 5 were smaller than for male students, indicated by negative scale score differences in panels for these math grades (Figure 4 and Table 4a). Scale score differences between these groups within these grades increased in 2022, compared to 2019, indicated by negative line slopes in the same panels. In grades 6 – 8, however, female students performed better in mathematics than their male counterparts in 2019, indicated by positive scale score differences for 2019 in panels for these grades. In 2022, however, scale score differences for female students in these grades, became slightly negative (in grades 6 and 7) and/or were almost zero (in grades 7 and 8). Put differently, female students performed slightly worse in grade 6 and 7 and only slightly better in grade 8 than male students in 2022.

In reading, female students performed consistently better than male students across all grades and academic years, indicated by positive scale score differences in favor of the female group (Figure 4 and Table 4a). Moreover, beginning in grade 4, scale score differences between these groups increased across grades. In other words, the performance gap in reading increased

between female and male students, beginning in grade 4. Within each grade, however, scale score differences between these groups decreased in grades 3, 7, and 8 (indicated by negative line slopes within panels for these grades), and increased in grades 4, 5, and 6 (indicated by positive line slopes within panels for these grades) in 2022.
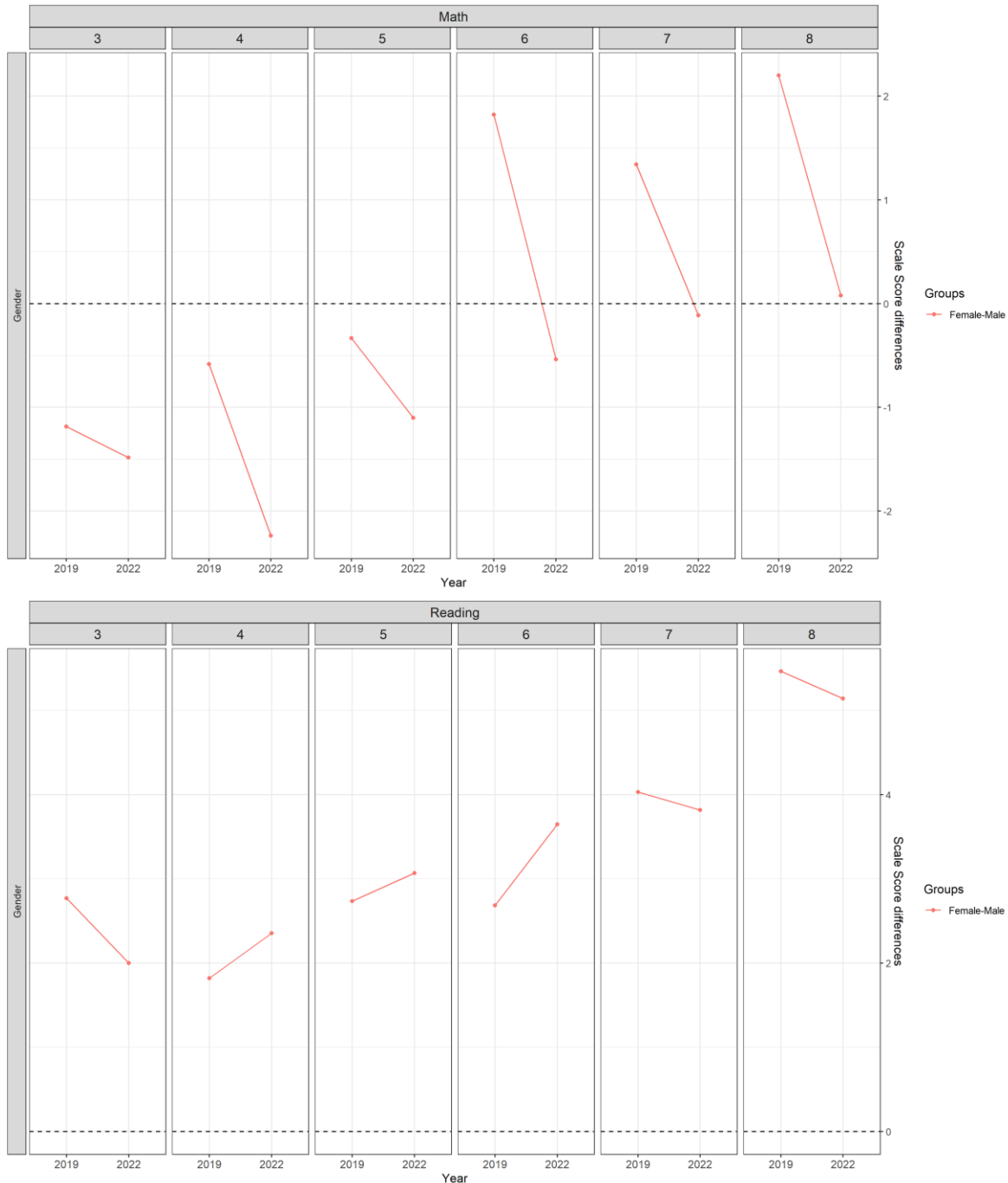
Figure 4. Math and reading scale score differences between groups based on gender across grades

We compared math and reading mean scale scores for Black and White students, Hispanic and White students, and Hispanic and Black students. The White group served as the reference group in comparisons which involved this group. Similarly, the Hispanic group served as the reference group in the comparison between Black and Hispanic students.

2019 and 2022 mathematics mean scale scores for Black and Hispanic students were consistently smaller than for White students across grades, indicated by negative scale score differences for these comparisons in all math panels (Figure 5 and Table 4b). Absolute scale score differences between Hispanic and White students were smaller than those between Black and White students across grades and across years. These observations indicated that White students performed consistently better in mathematics than Black and Hispanic students across grades and years, with smaller performance gaps between Hispanic and White students than between Black and White students. Similar observations were made for reading scale score differences between these groups.

In 2019, Hispanic students performed consistently better than Black students in mathematics across all grades, indicated by positive scale score differences in favor of the Hispanic group for this year. 2022 math scale score differences between these two groups decreased in grade 3 compared to 2019. Beginning in grade 4, however, 2022 math scale score differences between the two groups became negative (indicated by a switch in scale score differences across the reference line of zero in Figure 5). In other words, 2022 math performance of Hispanic students became worse than 2022 performance of Black students in these grades. Similar observations were made for reading scale score differences between these groups (Figure 5 and Table 4b).

2019 and 2022 scale score differences between Black and White students were generally stable within grades 3 – 5 (indicated by generally horizontal slope lines in these panels) and then decreased in grades 6 – 8 (indicated by positive slope lines in these panels). Similar observations were made for math scale differences between Hispanic and White students, with the exception of math grade 3, in which there was an increase in scale score differences in 2022. 2022 reading scale score differences between Black and White students and between Hispanic and White students, however, decreased in grade 3 (indicated by a positive line slope in the reading grade 3 panel), generally increased in grades 4 – 7 (indicated by negative line slopes in panels for these reading grades), and decreased again in grade 8 (Figure 5).
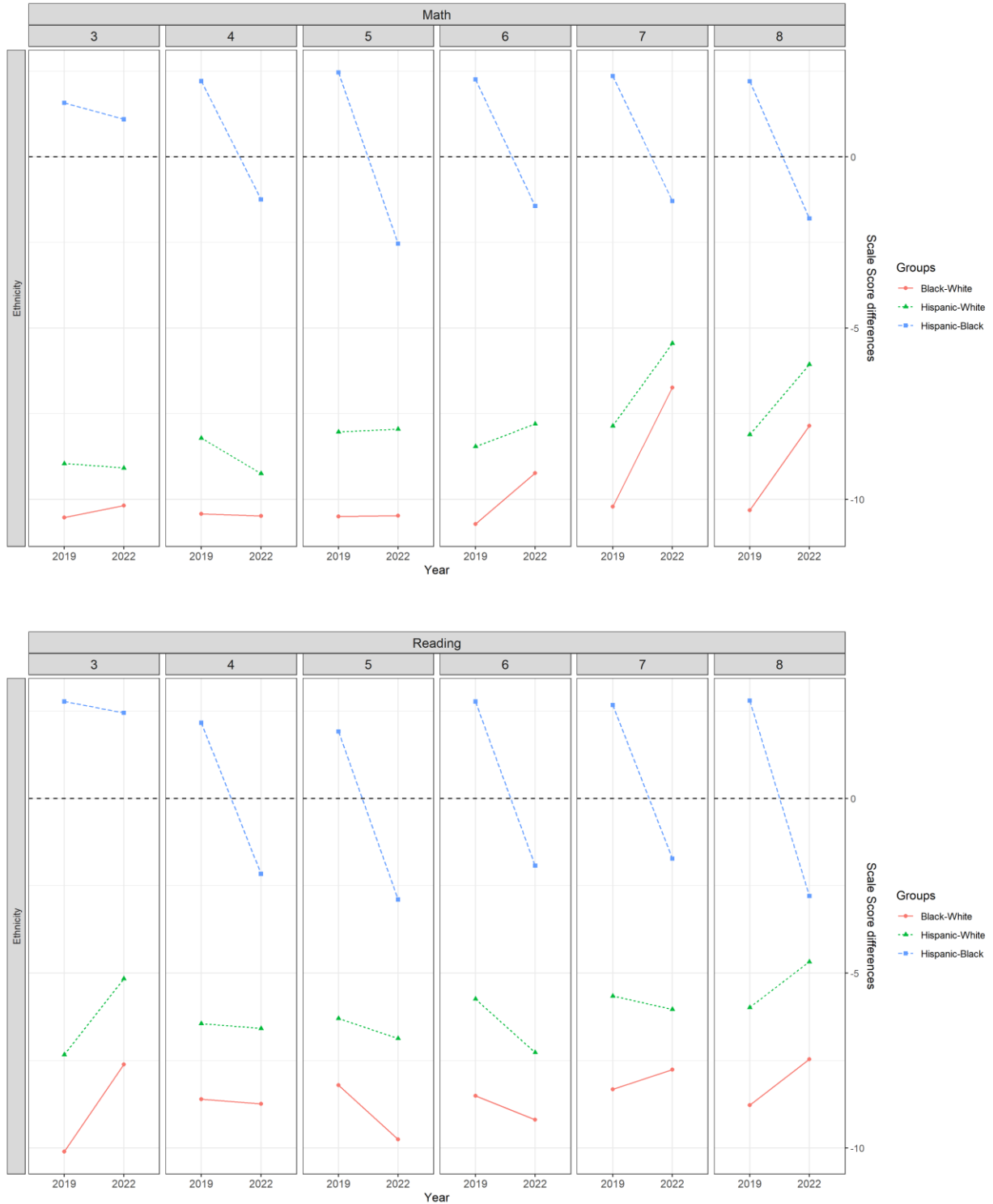
Figure 5. Math and reading scale score differences between groups based on ethnicity across grades

Comparisons based on English language proficiency

We compared mathematics and reading mean scale scores based on students' English language proficiency (a categorical variable with two groups, *Yes/No*, with *Yes* indicating limited English proficiency). Students who were English proficient served as the reference group in these comparisons.

2019 and 2022 mathematics mean scale scores for students with limited English proficiency (the *Yes* group) in all grades were smaller than scores for English proficient students, indicated by negative scale score differences in panels for these math grades (Figure 6 and Table 4c). So, students with limited English proficiency generally perform worse than English proficient students in mathematics across grades. These scale score differences generally got larger from grades 3 – 6, indicating that math performance gaps between these two groups increased as grades increased. Similar observations were made for reading scale score differences (Figure 6 and Table 4c).

Except for math grade 4, 2022 scale score differences between these groups within each grade generally decreased compared to 2019, indicated by positive line slopes in each of the panels. Larger decreases in scale score differences occurred in grades 6 – 8. Similar observations were made for reading scale score differences (Figure 6 and Table 4c).
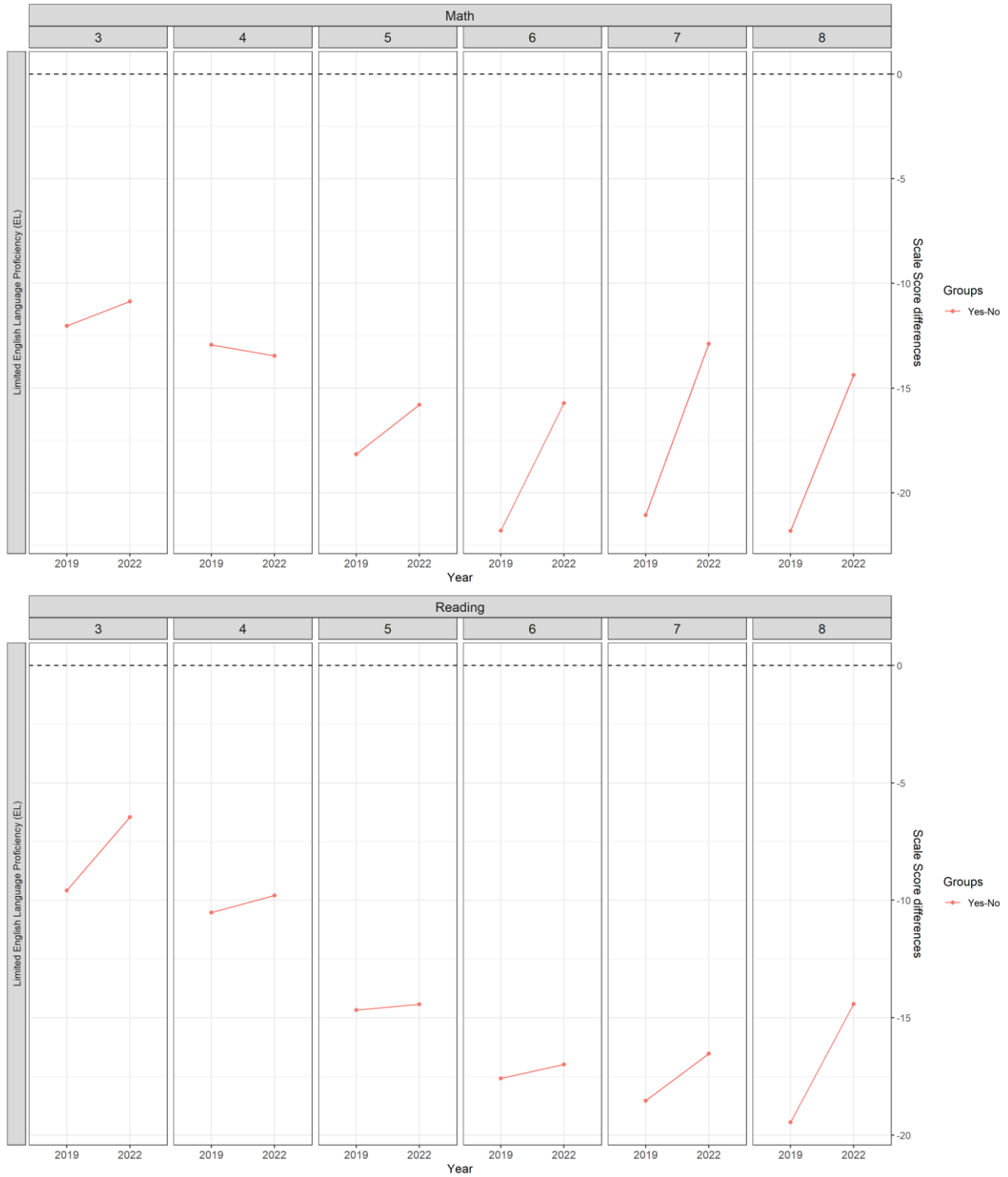
Figure 6. Math and reading scale score differences between groups based on English proficiency across grades

Comparisons based on individual education plan

We compared mathematics and reading mean scale scores based on whether students had or did not have individual education plans (IEP, a categorical variable with two groups, *Yes*/*No*, with *Yes* indicating students who have individual education plans). Students who did not have individual education plans served as the reference group in these comparisons.

2019 and 2022 mathematics mean scale scores for students with IEPs (the *Yes* group) in all grades were smaller than scores for students who did not have IEPs, indicated by negative scale score differences in panels for these math grades (Figure 7 and Table 4d). This indicated that students with IEPs generally performed worse than students who did not have IEPs in mathematics across grades. Moreover, math scale score differences increased slightly across grades, indicating that math performance gaps between these two groups increased as grades increased. Similar observations were made for reading scale score differences (Figure 7 and Table 4d).

With the exception of math grade 4, 2022 scale score differences between these groups within each grade generally decreased compared to 2019, indicated by positive line slopes in each of the panels. Larger decreases in scale score differences occurred in grades 6 – 8. Unlike mathematics, however, 2022 reading scale score differences in grades 4 – 7 generally increased as grades increased, indicated by negative line slopes in each of the panels (Figure 7). In other words, 2022 reading performance gaps widened in these grades.
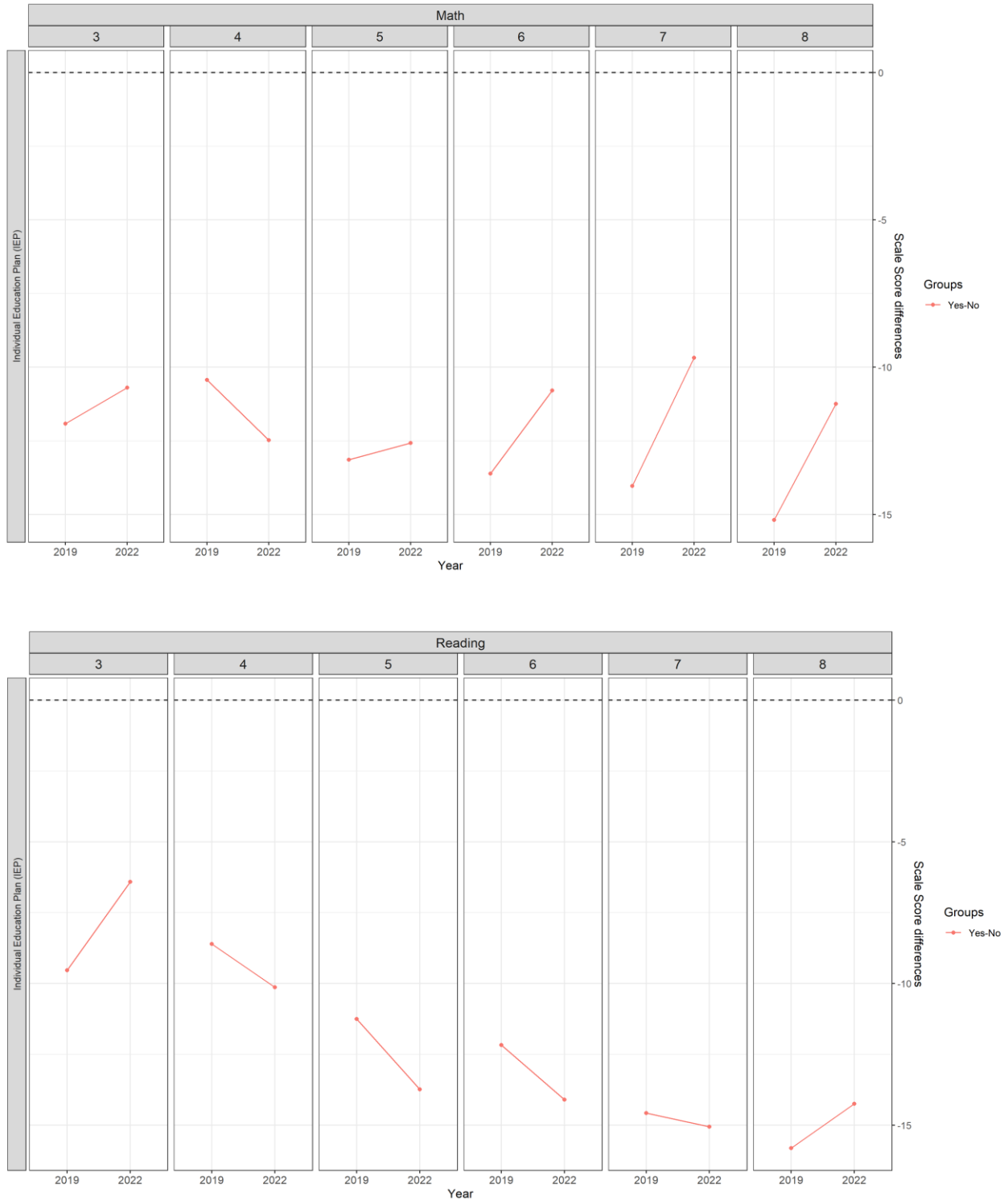
Figure 7. Math and reading scale score differences between groups based on education plan across grades

Comparisons based on eligibility for free/reduced Lunch

We compared math and reading mean scale scores for students who were eligible for free lunch (the *Free* lunch group) and students who were eligible for reduced lunch (the *Reduced* lunch group) with students who were not eligible for either (the *Not Eligible* group), as well as between the *Free* group and the *Reduced* group. Students who were not eligible for either free or reduced lunch served as the reference group in comparisons involving this group. Students in the *Reduced* lunch group served as the reference group in the comparison between the *Free* lunch group and this group.

2019 and 2022 mathematics and reading mean scale scores for the *Free* and *Reduced* groups were consistently smaller than for the *Not Eligible* group across grades, indicated by negative scale score differences for these comparisons in all math and reading panels (Figure 8 and Table 4e). Absolute scale score differences between the *Free* and *Not Eligible* group were smaller than those between the *Reduced* and *Not Eligible* groups across grades and across years. These observations indicated that students who were not eligible for free or reduced lunch performed consistently better in mathematics and reading than students who were eligible for free lunch and those who were eligible for reduced lunch. Smaller performance gaps existed between the latter groups than between the former groups. Similarly, 2019 and 2022 mathematics and reading mean scale scores for students who were eligible for free lunch were consistently smaller than for students who were eligible for reduced lunch across grades, indicated by negative scale score differences for these comparisons in all math and reading panels (Figure 8 and Table 4e).

2022 mathematics and reading mean scale score differences between the *Free* and the *Not Eligible* groups and between the *Reduced* and the *Not Eligible* groups generally increased across grades 3 – 6 (indicated by negative line slopes in corresponding panels in Figure 8) but decreased in grades 7 and 8 (positive line slopes in the corresponding panels in Figure 8). For the *Free* and *Reduced* groups, math mean scale score differences within each grade decreased in 2022. For the same groups, however, reading mean scale scores increased in grades 3 – 7.
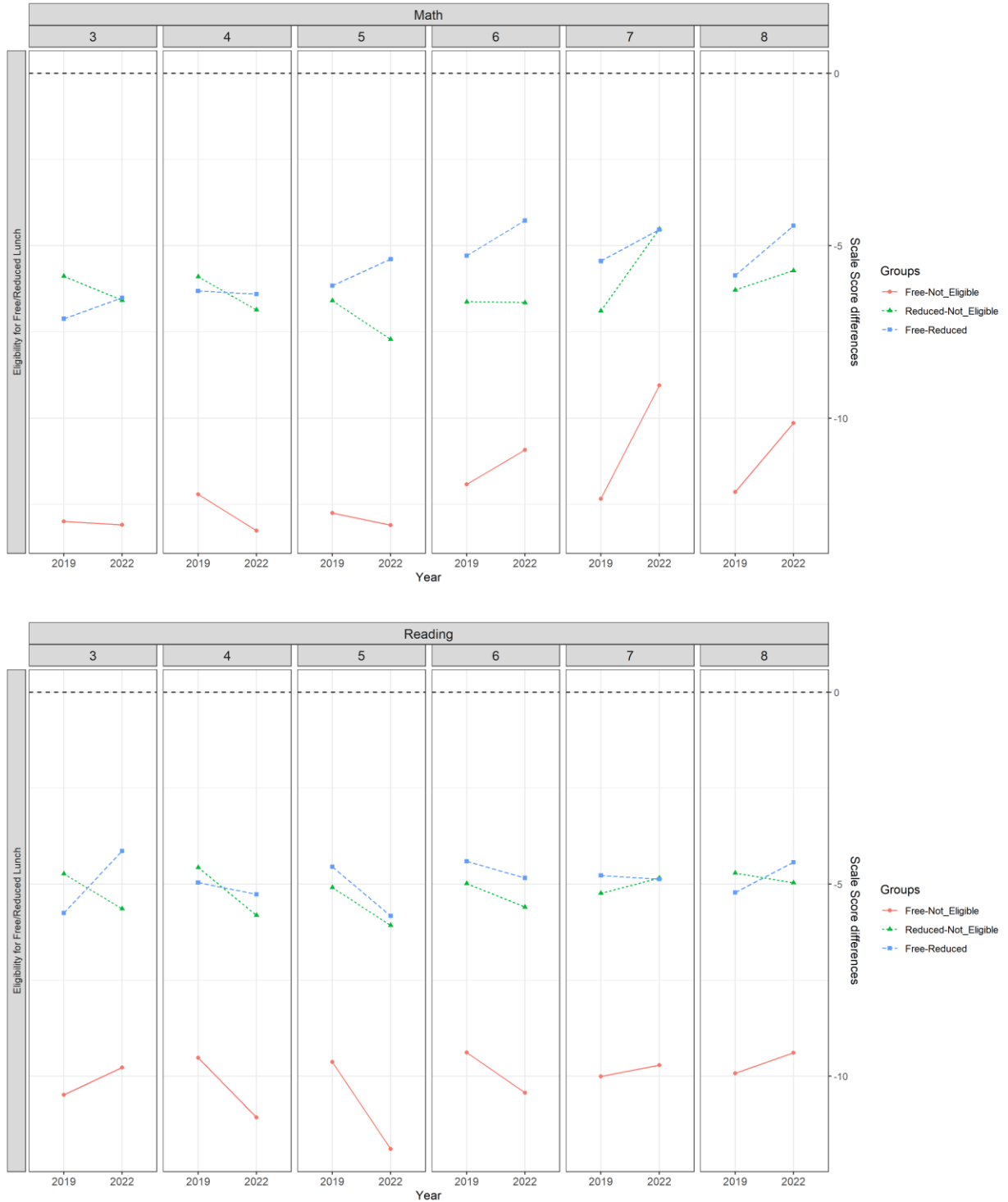
Figure 8. Math and reading scale score differences amongst groups based on eligibility for free/reduced lunch across grades

Within-group comparisons

We also compared 2022 mean scale scores to 2019 mean scale scores in mathematics and reading of students who belonged to the same groups identified by gender, ethnicity, English language proficiency, education plan, and eligibility for free/reduced lunch. For example, we compared performance of female students who took the 2022 operational assessments to female students who took the 2019 operational assessments. We conducted similar comparisons on performances of male students on the 2022 and 2019 operational assessments, as well as on performances for other groups identified by the variables described above. We examined differences within these groups within and across grades. Figures 9 through 13 depict these comparisons graphically. Tables 5a through 5e in Appendix A present the same information in tabular form.

Comparisons based on gender

In general, 2022 mean scale scores for female and male students were smaller than 2019 mean scale scores for the same groups, respectively, in both mathematics and reading. These are indicated by negative scale score differences in all math and reading panels in Figure 9 (and in Table 5a). 2022 – 2019 mean scale score differences for each group followed similar patterns of increases and decreases across grades. Group differences in each subject generally increased as grades increased.
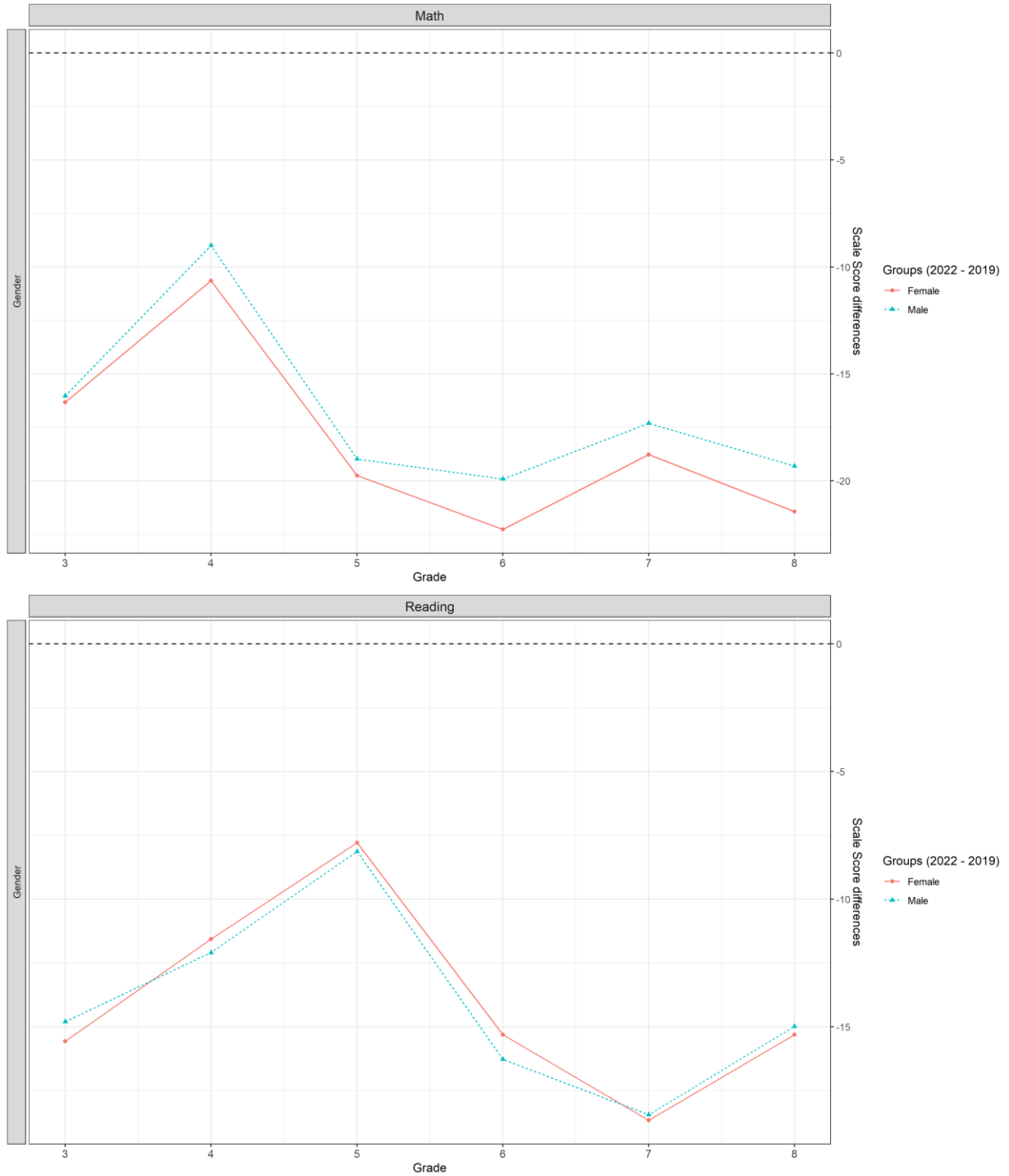
Figure 9. Math and reading scale score differences within groups based on gender across grades

Comparisons based on ethnicity

2022 mean scale scores for White, Black, and Hispanic students were smaller than 2019 mean scale scores for the same groups, respectively, in both mathematics and reading. These are indicated by negative scale score differences in all math and reading panels in Figure 10 (and in Table 5b). 2022 – 2019 mean scale score differences for each group followed similar patterns of increases and decreases across grades. Group differences in each subject generally increased as grades increased. For mathematics grades 6 – 8, 2022 – 2019 scale score differences for Black and Hispanic students were slightly smaller than differences for White students.
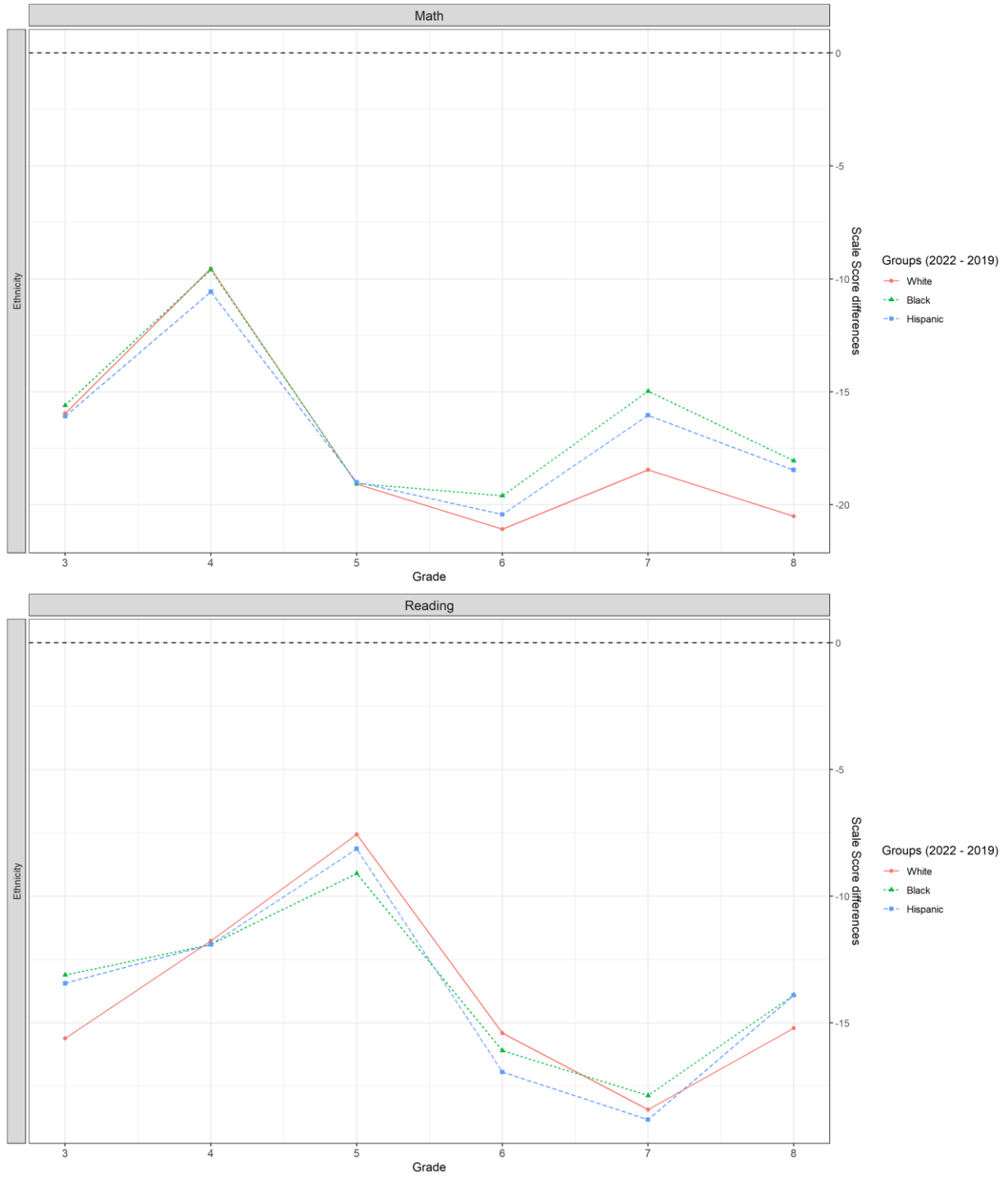
Figure 10. Math and reading scale score differences within groups based on ethnicity across grades

Comparisons based on English language proficiency

2022 mean scale scores for students with limited English proficiency (the *Yes* group) and students who were proficient in English (the *No* group) were smaller than 2019 mean scale scores for the same groups, respectively, in both mathematics and reading. These are indicated by negative scale score differences in all math and reading panels in Figure 11 (and in Table 5c). 2022 – 2019 mean scale score differences for students with limited English proficiency in mathematics grades 3 and 5 – 8 were smaller than differences for students who were English proficient. A similar observation was made for 2022 – 2019 mean scale score differences between both groups across all reading grades.
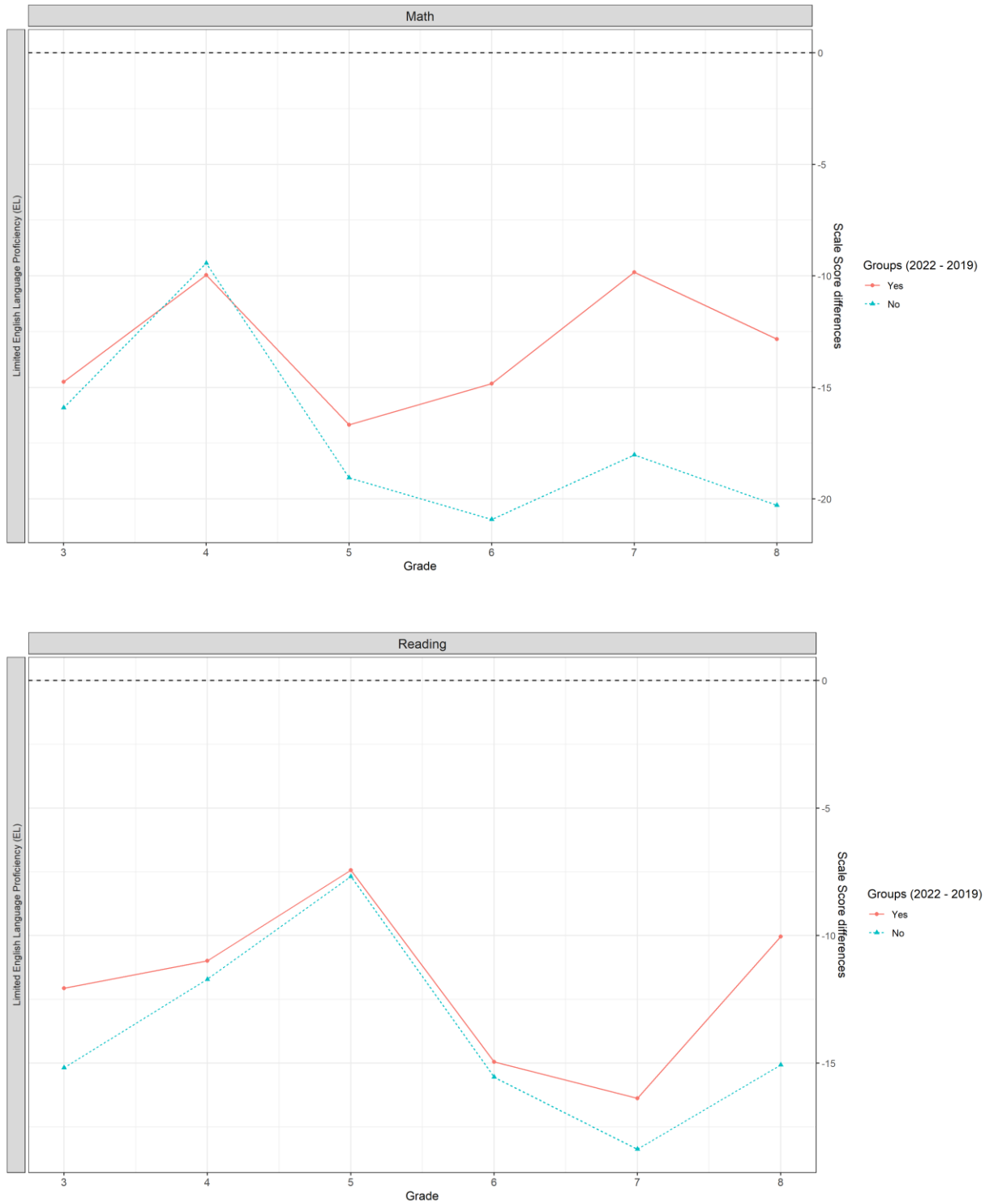
Figure 11. Math and reading scale score differences within groups based on English proficiency across grades

Comparisons based on individual education plans

2022 mean scale scores for students with individual education plans (the *Yes* group) and students who did not have IEPs (the *No* group) were smaller than 2019 mean scale scores for the same groups, respectively, in both mathematics and reading. These are indicated by negative scale score differences in all math and reading panels in Figure 12 (and in Table 5d). 2022 – 2019 mean scale score differences for students with IEPs in mathematics grades 3 and 5 – 8 were smaller than differences for students who did not have IEPs. In contrast to this observation for math grades, 2022 – 2019 mean scale score differences for students who did not have IEPs were smaller than differences for students with IEPs in reading grades 4 – 7.
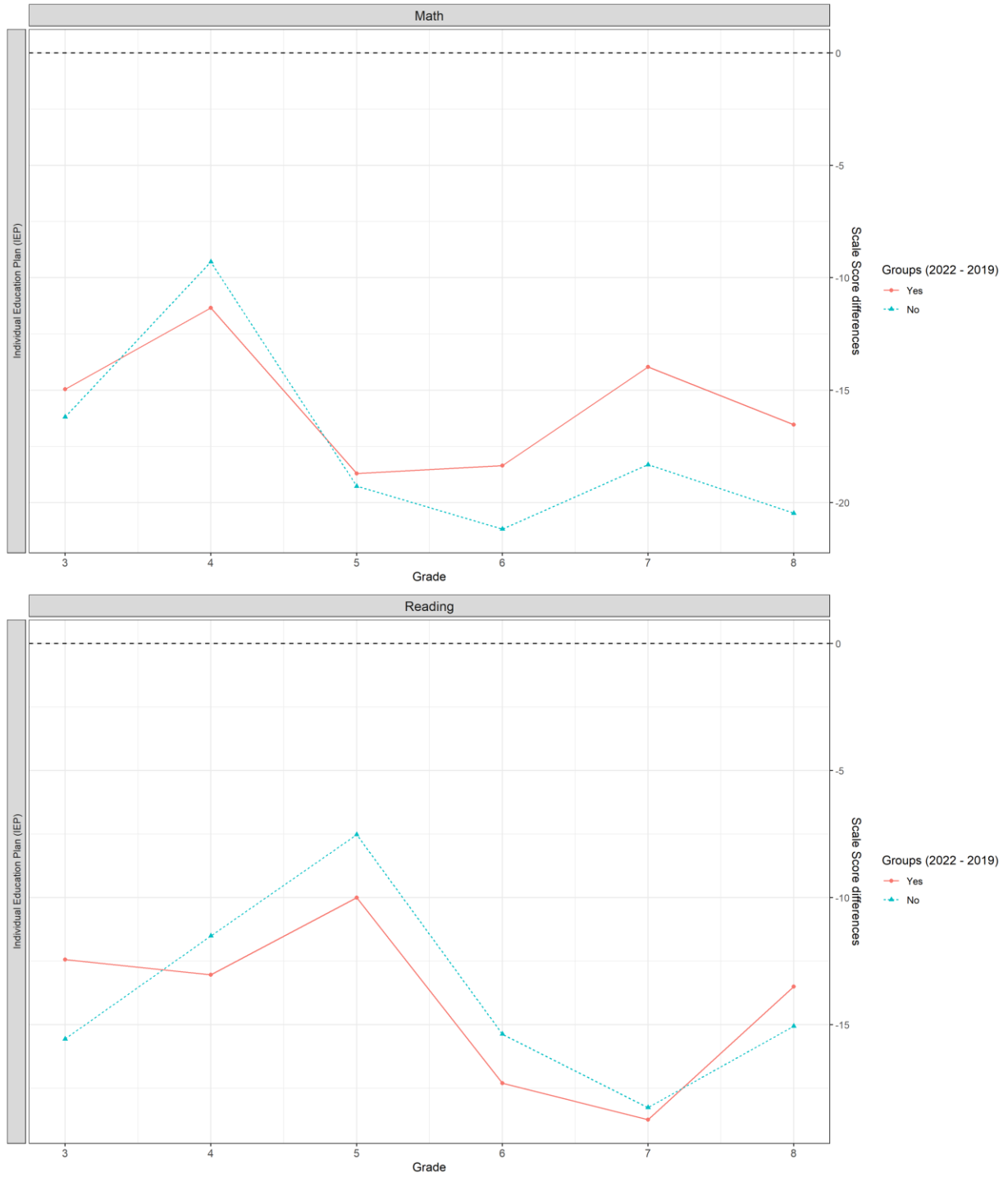
Figure 12. Math and reading scale score differences within groups based on education plan across grades

Eligibility for Free/Reduced Lunch (Lunch)

2022 mean scale scores for students who were eligible for free lunch (the *Free* lunch group), students who were eligible for reduced lunch (the *Reduced* lunch group), and students who were not eligible for either (the *Not Eligible* group) were smaller than 2019 mean scale scores for the same groups, respectively, in both mathematics and reading. These are indicated by negative scale score differences in all math and reading panels in Figure 11 (and in Table 5e). 2022 – 2019 mean scale score differences for each group followed similar patterns of increases and decreases across grades. Group differences in each subject generally increased as grades increased. 2022 – 2019 mean scale score differences for the *Free* lunch group and for the *Reduced* lunch group in mathematics grades 7 and 8 were slightly smaller than differences for the *Not Eligible* group.
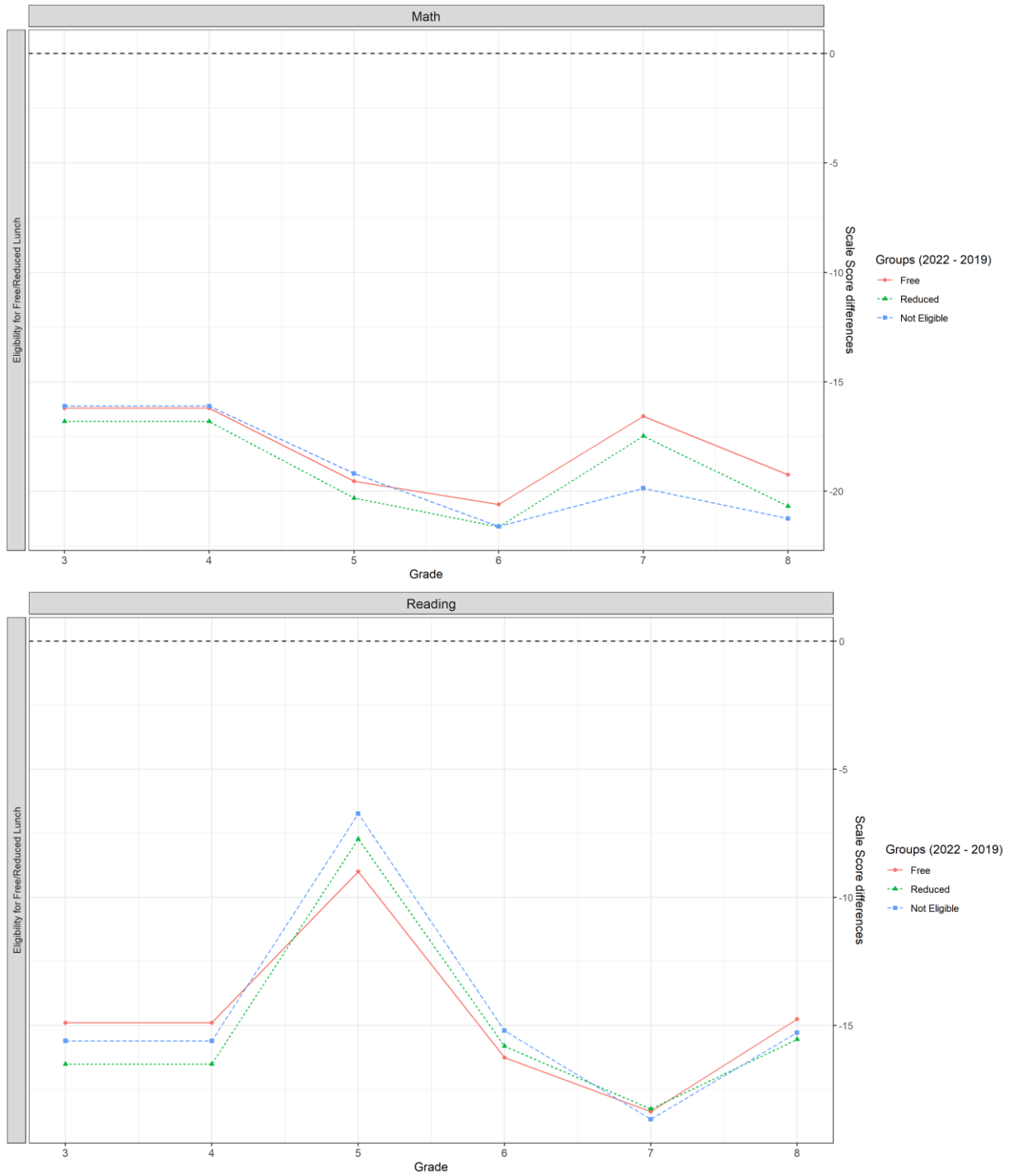
Figure 13. Math and reading scale score differences within groups based on eligibility for free/reduced lunch across grades

A Synopsis of Key Results on Student Performance on the Kentucky State Assessments (KSA) and on the National Assessment of Educational Progress (NAEP) 2022 Mathematics and Reading Grades 4 and 8 assessments

We compared results on students' performance on KSA with results from the National Assessment of Educational Progress (NAEP) 2022 assessments. We describe key findings from both sets of results below. A comprehensive description of student group performance results on NAEP 2022 mathematics and reading grades 4 and 8 assessments is provided in Appendix B.

Overall comparisons

Similar to performance on KSA, Kentucky students experienced overall decreases in NAEP mathematics and reading scores from 2019 to 2022. Students' NAEP mathematics and reading grade 4 scale scores declined by five points and four points (on the NAEP 0 – 500 scoring scale), respectively (Table 3 and NAEP math and reading grade 4 panels in Figure 14). Similarly, NAEP mathematics and reading grade 8 scale scores declined by nine and five points, respectively (Table 3 and NAEP math and reading grade 8 panels in Figure 14). These overall NAEP results are similar to results in this study in which we observed 10-point and 12-point overall declines (on the KSA scoring scales) on KSA mathematics and reading grade 4 assessments, respectively, and 20-point and 15-point declines on KSA math and reading grade 8 assessments, respectively (Table 2 and Figure 2). Both KSA and NAEP assessments for these subjects and grades indicate a decline in student performance from 2019 to 2022. These declines in math and reading performances are larger in the higher grade.

Table 3. Results for Kentucky from the National Assessment of Educational Progress (NAEP) 2022 mathematics and reading grades 4 and 8 assessments (Scale 0-500)

Variable	Group	G4 MA 2019	G4 MA 2022	G4 RD 2019	G4 RD 2022	G8 MA 2019	G8 MA 2022	G8 RD 2019	G8 RD 2022
Overall	All Students	239	234	221	217	278	269	263	258
Gender	Male	239	236	218	215	276	269	258	253
Gender	Female	239	232	224	218	280	269	268	264
Ethnicity	White	242	237	225	220	282	274	266	262
Ethnicity	Black	223	210	199	200	256	248	239	239
Ethnicity	Hispanic	228	227	206	207	270	259	258	254
Lunch	Lunch eligible	230	224	212	206	267	259	253	249
Lunch	Not lunch eligible	253	248	235	232	292	283	277	272

Note. G4= Grade 4. G8= Grade 8. MA= Mathematics. RD= Reading. Lunch= Free/reduced lunch. Data compiled from
https://www.nationsreportcard.gov/profiles/stateprofile/overview/KY?cti=PgTab_OT&chort=1&sub=MAT&sj=KY&fs=Grade&st=MN&year=2022R3&sg=Gender%3A%20Male%20vs.%20Female&sgv=Difference&ts=Single%20Year&tss=2022R3&sfj=NP
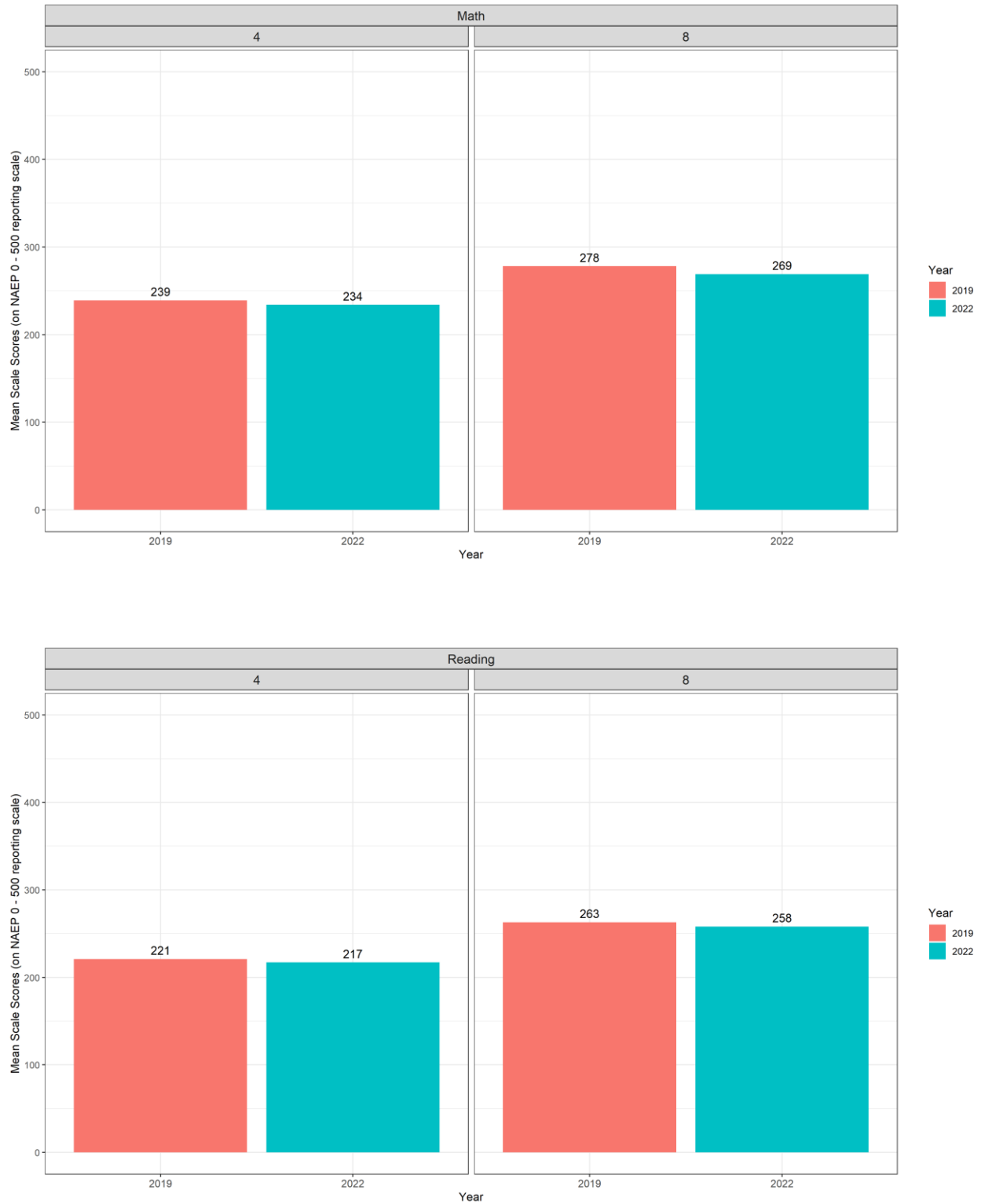
Figure 14. Overall results for Kentucky from NAEP 2022 Mathematics and reading grades 4 and 8 assessments

Comparisons by gender

Changes in group performance by gender from 2019 to 2022 are consistent on NAEP and KSA mathematics grades 4 and 8 assessments. Female students in grade 4 performed equally well as male students on both KSA and NAEP 2019 mathematics grade 4 assessments (KSA and NAEP math grade 4 panels in Figures 4 and 15, respectively). Female students in grade 8 performed better than male students on both KSA and NAEP 2019 mathematics grade 8 assessments (KSA and NAEP math grade 8 panels in Figures 4 and 15, respectively). While KSA 2022 and NAEP 2022 scale scores in both math grades declined for both groups, declines in scale scores for female students were larger than declines in scale scores for male students for both assessments and grades.

Female students' KSA and NAEP 2019 scale scores on reading grades 4 and 8 were larger than male students' scale scores on the same statements, indicating that female students in both grades performed better on these assessments (KSA and NAEP reading grades 4 and 8 panels in Figures 4 and 15, respectively). In contrast to larger scale score differences in KSA 2022 reading grade 4 differences between male and female students increased, with female students continuing to perform better than male students. NAEP 2022 scale score differences between gender groups for the same grade decreased. The opposite observation for reading grade 8 can be made: KSA 2022 reading grade 8 scale score differences between groups decreased, while differences increased on NAEP 2022.

Comparisons by ethnicity

Order of student performance by ethnicity group on KSA and NAEP 2019 mathematics and reading grades 4 and 8 were similar. Larger scale score differences in mathematics and in reading between Black and White students compared to differences between Hispanic and White students indicated lower performance by Black students than Hispanic students, compared to White students, in both subjects and grades in both assessments (KSA and NAEP math and reading grades 4 and 8 panels in Figures 5 and 15, respectively). This order of performance by ethnicity group is maintained in both grades on KSA and NAEP 2022 math and reading assessments. However, patterns of group performance in both grades from 2019 to 2022 were inconsistent on both assessments. For example, math grade 4 results on KSA and NAEP assessments indicated contrasting performances between Black and White students, between Hispanic and White students, and between Hispanic and Black students. On KSA, math grade 4 scale score differences from 2019 to 2022 between Black and White students were stable. However, NAEP results indicated an increase in scale score differences for the same group comparison. KSA 2022 results indicated an increase in scale score differences between Hispanic and White students from 2019 while NAEP 2022 results indicated a decrease. KSA math grade 4 performance of Hispanic students became worse than performance of Black students while NAEP indicated an increased difference in performance between the two groups, with larger scale scores for the Hispanic group. KSA 2022 reading grade 4 indicated slight increases in score differences between Hispanic and White students and between Black and White students while NAEP results for the same grade indicated decreases in score differences for both comparisons. Other similarly inconsistent observations on KSA and NAEP results can be made for other group comparisons in both subjects and grades.

Comparisons by eligibility for free lunch

Scale scores for students who were eligible for free lunch were consistently smaller than scale scores for ineligible students on KSA and NAEP 2019 math and reading grades 4 and 8 (KSA and NAEP math and reading grades 4 and 8 panels in Figures 8 and 17, respectively). This order of performance is maintained in both grades on KSA and NAEP 2022 math and reading assessments. Additionally, similar patterns of performance from 2019 to 2022 on both assessments were observed for subjects and grades. For example, scale score differences on KSA and NAEP math grade 4 increased from 2019 to 2022 while differences on both assessments in math grade 8 decreased during the same period. Similar observations for reading grades 4 and 8 were made.

Discussion

The current study provides evidence of the impact of the COVID-19 event on student academic performance on Kentucky State Assessments. Across subjects and grades, 2022 mean scale scores for students who took KSA operational tests were consistently smaller than 2019 mean scale scores. 2022 – 2019 mean math scale score differences generally increased as grade level increased. For reading, 2022 – 2019 mean scale score differences decreased in grades 3 – 5 and then increased in the higher grades. On average, all students experienced losses following COVID-19. The impacts on disaggregated student groups, however, were less systematic.

KSA 2019 and 2022 math scale score differences between female and male students were comparable across grades, with differences no larger than 3 points on the KSA scoring scale. KSA 2022 math scale scores were generally smaller for female students than for male students while KSA 2022 reading scale scores were generally smaller for male students than for female students. KSA 2019 and 2022 mathematics and reading mean scale scores in all grades for students with limited English proficiency were consistently smaller than scores for English proficient students. Reading scale score differences between these groups generally increased as grade levels increased. Similar observations for comparisons based on groups identified by individual education plans were made. KSA 2019 and 2022 mathematics and reading mean scale scores for students who were eligible for free lunch (the *Free* lunch group) and for students who were eligible for reduced lunch (the *Reduced* lunch group) were consistently smaller than for the *Not Eligible* group across grades.

Of the scale scores for Black, Hispanic, and White student groups that were examined in this study, 2019 and 2022 KSA math and reading scale scores for Black students were consistently the smallest. This was followed by scale scores for Hispanic students across both subjects and grades. Math scale score differences between Black and White students and between Hispanic and White students were generally stable from 2019 to 2022 in grades 3 – 5 but these differences decreased in grades 6 – 7. Reading scale score differences for the same group comparisons were less systematic, with grade-specific fluctuations in these differences. 2022 scale scores for Hispanic students were larger than scores for Black students in mathematics and reading grade 3 but became smaller in higher grades. Put differently, 2022 math and reading performance of Hispanic students became worse than 2022 performance of Black students.

NAEP 2019 and 2022 math and reading grades 4 and 8 results confirm observations regarding overall student performance. Student performance on KSA 2022 math and reading declined from 2019. However, patterns of score differences from 2019 to 2022 for groups identified by gender, ethnicity, EL proficiency, IEP, and eligibility for free or reduced lunch were subject- and grade-specific rather than systematic across subjects and grades.

References

Bynum, B. H., & Thacker, A. A. (2020). *Third-party checking of 2019 scaling and equating for the Kentucky Performance Rating for Educational Progress (K-PREP) Tests* (2020 No. 009). Human Resources Research Organization.

Thacker, A. A. (2021). *Kentucky's response to the COVID-19 pandemic: Steps toward ensuring continued assessment program relevance and validity of student scores* (2021 No. 073). Human Resources Research Organization.

Table 4a. Between-groups comparisons based on gender

Grade	Year	Comparisons	Math Scale Score differences	Math 95% Confidence Interval	Reading Scale Score differences	Reading 95% Confidence Interval
3	2019	Female – Male	-1.19	[-1.67, -0.70]	2.77	[2.35, 3.18]
3	2022	Female – Male	-1.48	[-1.97, -0.99]	2.00	[1.58, 2.43]
4	2019	Female – Male	-0.58	[-1.03, -0.14]	1.82	[1.44, 2.20]
4	2022	Female – Male	-2.24	[-2.72, -1.76]	2.35	[1.95, 2.76]
5	2019	Female – Male	-0.33	[-0.78, 0.12]	2.73	[2.34, 3.13]
5	2022	Female – Male	-1.10	[-1.57, -0.63]	3.07	[2.66, 3.48]
6	2019	Female – Male	1.82	[1.43, 2.22]	2.68	[2.32, 3.05]
6	2022	Female – Male	-0.54	[-0.95, -0.13]	3.65	[3.27, 4.03]
7	2019	Female – Male	1.34	[0.97, 1.72]	4.03	[3.67, 4.40]
7	2022	Female – Male	-0.11	[-0.49, 0.27]	3.82	[3.45, 4.19]
8	2019	Female – Male	2.20	[1.80, 2.60]	5.47	[5.11, 5.82]
8	2022	Female – Male	0.08	[-0.31, 0.47]	5.14	[4.79, 5.49]

Table 4b. Between-groups comparisons based on ethnicity

Grade	Year	Comparisons	Math Scale Score differences	Math 95% CI	Reading Scale Score differences	Reading 95% CI
3	2019	Black - White	-10.54	[-11.26, -9.81]	-10.11	[-10.74, -9.47]
3	2019	Hispanic – White	-8.96	[-9.93, -7.99]	-7.33	[-8.19, -6.48]
3	2019	Hispanic – Black	1.58	[0.44, 2.72]	2.77	[1.77, 3.77]
3	2022	Black - White	-10.19	[-10.92, -9.45]	-7.61	[-8.25, -6.98]
3	2022	Hispanic – White	-9.09	[-10.05, -8.13]	-5.16	[-6.01, -4.32]
3	2022	Hispanic – Black	1.10	[-0.03, 2.22]	2.45	[1.46, 3.44]
4	2019	Black - White	-10.43	[-11.1, -9.76]	-8.61	[-9.18, -8.04]
4	2019	Hispanic – White	-8.22	[-9.12, -7.31]	-6.44	[-7.22, -5.67]
4	2019	Hispanic – Black	2.21	[1.16, 3.26]	2.16	[1.26, 3.07]
4	2022	Black - White	-10.49	[-11.22, -9.77]	-8.74	[-9.36, -8.12]
4	2022	Hispanic – White	-9.25	[-10.19, -8.31]	-6.58	[-7.4, -5.77]
4	2022	Hispanic – Black	-1.24	[-2.35, -0.13]	-2.16	[-3.11, -1.20]
5	2019	Black - White	-10.51	[-11.18, -9.83]	-8.21	[-8.80, -7.61]
5	2019	Hispanic – White	-8.04	[-8.93, -7.14]	-6.29	[-7.09, -5.50]
5	2019	Hispanic – Black	2.47	[1.42, 3.52]	1.91	[0.98, 2.85]
5	2022	Black - White	-10.49	[-11.18, -9.79]	-9.76	[-10.38, -9.14]
5	2022	Hispanic – White	-7.95	[-8.85, -7.05]	-6.87	[-7.67, -6.06]
5	2022	Hispanic – Black	-2.53	[-3.59, -1.48]	-2.89	[-3.84, -1.95]
6	2019	Black - White	-10.72	[-11.32, -10.13]	-8.51	[-9.07, -7.96]
6	2019	Hispanic – White	-8.46	[-9.26, -7.66]	-5.74	[-6.50, -4.98]
6	2019	Hispanic – Black	2.26	[1.33, 3.20]	2.77	[1.89, 3.66]
6	2022	Black - White	-9.24	[-9.85, -8.63]	-9.19	[-9.77, -8.62]
6	2022	Hispanic – White	-7.80	[-8.60, -7.01]	-7.27	[-8.02, -6.52]
6	2022	Hispanic – Black	-1.44	[-2.37, -0.50]	-1.92	[-2.80, -1.05]
7	2019	Black - White	-10.22	[-10.80, -9.64]	-8.33	[-8.90, -7.75]
7	2019	Hispanic – White	-7.86	[-8.63, -7.09]	-5.65	[-6.42, -4.88]
7	2019	Hispanic – Black	2.36	[1.45, 3.27]	2.67	[1.77, 3.58]
7	2022	Black - White	-6.74	[-7.30, -6.18]	-7.76	[-8.32, -7.21]
7	2022	Hispanic – White	-5.45	[-6.18, -4.72]	-6.04	[-6.77, -5.31]
7	2022	Hispanic – Black	-1.29	[-2.15, -0.43]	-1.72	[-2.58, -0.86]
8	2019	Black - White	-10.32	[-10.93, -9.71]	-8.78	[-9.35, -8.21]
8	2019	Hispanic – White	-8.11	[-8.96, -7.26]	-5.98	[-6.77, -5.19]
8	2019	Hispanic – Black	2.21	[1.22, 3.20]	2.80	[1.88, 3.72]
8	2022	Black - White	-7.86	[-8.45, -7.27]	-7.47	[-8.01, -6.93]
8	2022	Hispanic – White	-6.06	[-6.82, -5.31]	-4.67	[-5.38, -3.97]
8	2022	Hispanic – Black	-1.80	[-2.69, -0.91]	-2.79	[-3.62, -1.96]

Table 4c. Between-groups comparisons based on limited English language proficiency

Grade	Year	Comparisons	Math Scale Score differences	Math 95% CI	Reading Scale Score differences	Reading 95% CI
3	2019	Yes - No	-12.03	[-12.99, -11.08]	-9.58	[-10.42, -8.73]
3	2022	Yes - No	-10.86	[-11.74, -9.99]	-6.46	[-7.24, -5.69]
4	2019	Yes - No	-12.93	[-13.87, -12.00]	-10.52	[-11.33, -9.71]
4	2022	Yes - No	-13.46	[-14.36, -12.55]	-9.79	[-10.58, -9.01]
5	2019	Yes - No	-18.16	[-19.26, -17.05]	-14.67	[-15.68, -13.66]
5	2022	Yes - No	-15.79	[-16.76, -14.83]	-14.42	[-15.3, -13.55]
6	2019	Yes - No	-21.80	[-22.92, -20.68]	-17.58	[-18.68, -16.49]
6	2022	Yes - No	-15.72	[-16.69, -14.74]	-16.99	[-17.92, -16.06]
7	2019	Yes - No	-21.07	[-22.24, -19.9]	-18.53	[-19.73, -17.33]
7	2022	Yes - No	-12.88	[-13.89, -11.88]	-16.53	[-17.55, -15.5]
8	2019	Yes - No	-21.81	[-22.99, -20.64]	-19.45	[-20.58, -18.32]
8	2022	Yes - No	-14.37	[-15.38, -13.37]	-14.41	[-15.37, -13.45]

Table 4d. Between-groups comparisons based on individual education plan

Grade	Year	Comparisons	Math Scale Score differences	Math 95% CI	Reading Scale Score differences	Reading 95% CI
3	2019	Yes - No	-11.92	[-12.57, -11.27]	-9.53	[-10.09, -8.97]
3	2022	Yes - No	-10.69	[-11.35, -10.04]	-6.41	[-6.98, -5.84]
4	2019	Yes - No	-10.43	[-11.05, -9.82]	-8.61	[-9.12, -8.09]
4	2022	Yes - No	-12.48	[-13.13, -11.82]	-10.13	[-10.68, -9.58]
5	2019	Yes - No	-13.14	[-13.77, -12.51]	-11.25	[-11.8, -10.7]
5	2022	Yes - No	-12.58	[-13.23, -11.92]	-13.73	[-14.3, -13.17]
6	2019	Yes - No	-13.61	[-14.20, -13.03]	-12.17	[-12.70, -11.63]
6	2022	Yes - No	-10.79	[-11.38, -10.20]	-14.10	[-14.63, -13.56]
7	2019	Yes - No	-14.03	[-14.61, -13.45]	-14.57	[-15.12, -14.02]
7	2022	Yes - No	-9.68	[-10.24, -9.12]	-15.05	[-15.58, -14.52]
8	2019	Yes - No	-15.18	[-15.82, -14.55]	-15.80	[-16.37, -15.24]
8	2022	Yes - No	-11.24	[-11.84, -10.65]	-14.24	[-14.77, -13.71]

Table 4e. Between-groups comparisons based on eligibility for free/reduced lunch

Grade	Year	Comparisons	Math Scale Score differences	Math 95% CI	Reading Scale Score differences	Reading 95% CI
3	2019	Free – Not Eligible	-13.00	[-13.53, -12.46]	-10.49	[-10.95, -10.02]
3	2019	Reduced - Not Eligible	-5.88	[-7.21, -4.55]	-4.73	[-5.89, -3.58]
3	2019	Free – Reduced	-7.12	[-8.42, -5.81]	-5.75	[-6.89, -4.62]
3	2022	Free – Not Eligible	-13.09	[-13.63, -12.54]	-9.78	[-10.25, -9.30]
3	2022	Reduced - Not Eligible	-6.58	[-8.04, -5.12]	-5.64	[-6.91, -4.37]
3	2022	Free – Reduced	-6.51	[-7.95, -5.07]	-4.14	[-5.38, -2.89]
4	2019	Free – Not Eligible	-12.21	[-12.70, -11.72]	-9.52	[-9.94, -9.11]
4	2019	Reduced - Not Eligible	-5.90	[-7.13, -4.67]	-4.57	[-5.61, -3.53]
4	2019	Free – Reduced	-6.31	[-7.52, -5.11]	-4.96	[-5.98, -3.94]
4	2022	Free – Not Eligible	-13.26	[-13.79, -12.73]	-11.07	[-11.52, -10.63]
4	2022	Reduced - Not Eligible	-6.86	[-8.30, -5.41]	-5.81	[-7.04, -4.59]
4	2022	Free – Reduced	-6.40	[-7.83, -4.97]	-5.26	[-6.47, -4.06]
5	2019	Free – Not Eligible	-12.75	[-13.24, -12.25]	-9.63	[-10.07, -9.20]
5	2019	Reduced - Not Eligible	-6.59	[-7.84, -5.34]	-5.08	[-6.19, -3.98]
5	2019	Free – Reduced	-6.16	[-7.39, -4.93]	-4.55	[-5.63, -3.46]
5	2022	Free – Not Eligible	-13.10	[-13.61, -12.58]	-11.90	[-12.35, -11.44]
5	2022	Reduced - Not Eligible	-7.72	[-9.09, -6.34]	-6.07	[-7.28, -4.86]
5	2022	Free – Reduced	-5.38	[-6.74, -4.03]	-5.82	[-7.02, -4.63]
6	2019	Free – Not Eligible	-11.91	[-12.35, -11.48]	-9.39	[-9.79, -8.98]
6	2019	Reduced - Not Eligible	-6.63	[-7.70, -5.55]	-4.98	[-5.98, -3.99]
6	2019	Free – Reduced	-5.29	[-6.34, -4.23]	-4.40	[-5.39, -3.42]
6	2022	Free – Not Eligible	-10.92	[-11.37, -10.47]	-10.43	[-10.85, -10.01]
6	2022	Reduced - Not Eligible	-6.65	[-7.86, -5.43]	-5.60	[-6.73, -4.46]
6	2022	Free – Reduced	-4.27	[-5.47, -3.07]	-4.84	[-5.95, -3.72]
7	2019	Free – Not Eligible	-12.34	[-12.75, -11.93]	-10.01	[-10.41, -9.61]
7	2019	Reduced - Not Eligible	-6.89	[-7.91, -5.87]	-5.24	[-6.24, -4.24]
7	2019	Free – Reduced	-5.44	[-6.45, -4.44]	-4.77	[-5.76, -3.79]
7	2022	Free – Not Eligible	-9.05	[-9.46, -8.64]	-9.71	[-10.12, -9.31]
7	2022	Reduced - Not Eligible	-4.51	[-5.64, -3.38]	-4.84	[-5.95, -3.73]
7	2022	Free – Reduced	-4.53	[-5.65, -3.42]	-4.87	[-5.97, -3.77]
8	2019	Free – Not Eligible	-12.14	[-12.57, -11.71]	-9.93	[-10.32, -9.53]
8	2019	Reduced - Not Eligible	-6.28	[-7.38, -5.18]	-4.71	[-5.71, -3.71]
8	2019	Free – Reduced	-5.86	[-6.95, -4.77]	-5.22	[-6.21, -4.23]
8	2022	Free – Not Eligible	-10.14	[-10.57, -9.71]	-9.39	[-9.78, -9.00]
8	2022	Reduced - Not Eligible	-5.72	[-6.91, -4.53]	-4.96	[-6.05, -3.87]
8	2022	Free – Reduced	-4.42	[-5.60, -3.24]	-4.43	[-5.51, -3.35]

Table 5a. Within-group comparisons based on gender

Grade	Comparison Years	Group	Math Scale Score differences	Math 95% CI	Reading Scale Score differences	Reading 95% CI
3	2022-2019	Female	-16.33	[-16.82, -15.83]	-15.56	[-15.99, -15.13]
3	2022-2019	Male	-16.03	[-16.51, -15.55]	-14.80	[-15.21, -14.38]
4	2022-2019	Female	-10.65	[-11.12, -10.18]	-11.56	[-11.96, -11.16]
4	2022-2019	Male	-9.00	[-9.45, -8.54]	-12.09	[-12.48, -11.71]
5	2022-2019	Female	-19.75	[-20.22, -19.29]	-7.79	[-8.20, -7.39]
5	2022-2019	Male	-18.98	[-19.44, -18.53]	-8.13	[-8.52, -7.73]
6	2022-2019	Female	-22.27	[-22.68, -21.86]	-15.31	[-15.69, -14.93]
6	2022-2019	Male	-19.91	[-20.31, -19.51]	-16.27	[-16.64, -15.90]
7	2022-2019	Female	-18.77	[-19.15, -18.39]	-18.66	[-19.03, -18.29]
7	2022-2019	Male	-17.31	[-17.68, -16.94]	-18.44	[-18.81, -18.08]
8	2022-2019	Female	-21.43	[-21.83, -21.03]	-15.31	[-15.67, -14.95]
8	2022-2019	Male	-19.31	[-19.70, -18.91]	-14.98	[-15.33, -14.63]

Table 5b. Within-group comparisons based on ethnicity

Grade	Comparison Years	Group	Math Scale Score differences	Math 95% CI	Reading Scale Score differences	Reading 95% CI
3	2022-2019	White	-15.95	[-16.38, -15.52]	-15.61	[-15.98, -15.23]
3	2022-2019	Black	-15.60	[-16.54, -14.66]	-13.11	[-13.93, -12.30]
3	2022-2019	Hispanic	-16.08	[-17.38, -14.78]	-13.44	[-14.58, -12.29]
4	2022-2019	White	-9.54	[-9.94, -9.13]	-11.77	[-12.12, -11.42]
4	2022-2019	Black	-9.60	[-10.50, -8.70]	-11.91	[-12.67, -11.14]
4	2022-2019	Hispanic	-10.57	[-11.81, -9.33]	-11.91	[-12.98, -10.84]
5	2022-2019	White	-19.08	[-19.49, -18.68]	-7.56	[-7.92, -7.20]
5	2022-2019	Black	-19.06	[-19.95, -18.18]	-9.11	[-9.89, -8.33]
5	2022-2019	Hispanic	-19.00	[-20.20, -17.79]	-8.13	[-9.21, -7.05]
6	2022-2019	White	-21.08	[-21.44, -20.73]	-15.41	[-15.75, -15.08]
6	2022-2019	Black	-19.60	[-20.37, -18.83]	-16.10	[-16.82, -15.37]
6	2022-2019	Hispanic	-20.43	[-21.50, -19.36]	-16.94	[-17.96, -15.93]
7	2022-2019	White	-18.45	[-18.78, -18.12]	-18.43	[-18.76, -18.11]
7	2022-2019	Black	-14.97	[-15.70, -14.23]	-17.87	[-18.60, -17.14]
7	2022-2019	Hispanic	-16.04	[-17.05, -15.03]	-18.82	[-19.83, -17.81]
8	2022-2019	White	-20.51	[-20.86, -20.16]	-15.21	[-15.53, -14.89]
8	2022-2019	Black	-18.05	[-18.82, -17.27]	-13.90	[-14.61, -13.18]
	2022-2019	Hispanic	-18.46	[-19.54, -17.38]	-13.91	[-14.92, -12.89]

Table 5c. Within-group comparisons based on limited English language proficiency

Grade	Comparison Years	Group	Math Scale Score differences	Math 95% CI	Reading Scale Score differences	Reading 95% CI
3	2022-2019	No	-15.91	[-16.27, -15.56]	-15.18	[-15.48, -14.87]
3	2022-2019	Yes	-14.75	[-15.99, -13.50]	-12.06	[-13.17, -10.96]
4	2022-2019	No	-9.43	[-9.77, -9.10]	-11.71	[-11.99, -11.43]
4	2022-2019	Yes	-9.96	[-11.21, -8.70]	-10.99	[-12.08, -9.89]
5	2022-2019	No	-19.05	[-19.37, -18.72]	-7.68	[-7.97, -7.40]
5	2022-2019	Yes	-16.68	[-18.11, -15.25]	-7.43	[-8.74, -6.13]
6	2022-2019	No	-20.92	[-21.20, -20.63]	-15.54	[-15.81, -15.28]
6	2022-2019	Yes	-14.83	[-16.29, -13.37]	-14.95	[-16.36, -13.53]
7	2022-2019	No	-18.02	[-18.31, -17.74]	-18.37	[-18.64, -18.09]
7	2022-2019	Yes	-9.84	[-11.35, -8.32]	-16.37	[-17.92, -14.81]
8	2022-2019	No	-20.28	[-20.56, -20.00]	-15.07	[-15.33, -14.82]
8	2022-2019	Yes	-12.84	[-14.36, -11.32]	-10.04	[-11.50, -8.58]

Table 5d. Within-group comparisons based on individual education plans

Grade	Comparison Years	Group	Math Scale Score differences	Math 95% CI	Reading Scale Score differences	Reading 95% CI
3	2022-2019	No	-16.19	[-16.56, -15.82]	-15.56	[-15.88, -15.24]
3	2022-2019	Yes	-14.96	[-15.81, -14.12]	-12.44	[-13.17, -11.71]
4	2022-2019	No	-9.29	[-9.64, -8.94]	-11.51	[-11.81, -11.22]
4	2022-2019	Yes	-11.34	[-12.16, -10.51]	-13.04	[-13.73, -12.34]
5	2022-2019	No	-19.27	[-19.61, -18.93]	-7.52	[-7.82, -7.22]
5	2022-2019	Yes	-18.70	[-19.55, -17.86]	-10.00	[-10.74, -9.27]
6	2022-2019	No	-21.18	[-21.48, -20.88]	-15.37	[-15.64, -15.10]
6	2022-2019	Yes	-18.36	[-19.13, -17.58]	-17.30	[-18.00, -16.59]
7	2022-2019	No	-18.31	[-18.59, -18.04]	-18.26	[-18.52, -17.99]
7	2022-2019	Yes	-13.96	[-14.72, -13.21]	-18.74	[-19.46, -18.02]
8	2022-2019	No	-20.47	[-20.76, -20.18]	-15.06	[-15.32, -14.80]
8	2022-2019	Yes	-16.53	[-17.35, -15.71]	-13.50	[-14.23, -12.76]

Table 5e. Within-group comparisons based on eligibility for free/reduced lunch

Grade	Comparison Years	Group	Math Scale Score differences	Math 95% CI	Reading Scale Score differences	Reading 95% CI
3	2022-2019	Free	-16.20	[-16.68, -15.73]	-14.89	[-15.30, -14.48]
3	2022-2019	Reduced	-16.81	[-18.70, -14.93]	-16.51	[-18.14, -14.87]
3	2022-2019	Not Eligible	-16.11	[-16.71, -15.51]	-15.60	[-16.12, -15.08]
4	2022-2019	Free	-16.20	[-16.68, -15.73]	-14.89	[-15.30, -14.48]
4	2022-2019	Reduced	-16.81	[-18.70, -14.93]	-16.51	[-18.14, -14.87]
4	2022-2019	Not Eligible	-16.11	[-16.71, -15.51]	-15.60	[-16.12, -15.08]
5	2022-2019	Free	-19.54	[-19.99, -19.09]	-8.99	[-9.39, -8.60]
5	2022-2019	Reduced	-20.31	[-22.09, -18.54]	-7.72	[-9.28, -6.15]
5	2022-2019	Not Eligible	-19.19	[-19.74, -18.63]	-6.73	[-7.22, -6.24]
6	2022-2019	Free	-20.61	[-21.01, -20.21]	-16.25	[-16.62, -15.88]
6	2022-2019	Reduced	-21.63	[-23.18, -20.08]	-15.81	[-17.25, -14.37]
6	2022-2019	Not Eligible	-21.61	[-22.09, -21.12]	-15.20	[-15.65, -14.75]
7	2022-2019	Free	-16.57	[-16.95, -16.20]	-18.36	[-18.72, -17.99]
7	2022-2019	Reduced	-17.48	[-18.94, -16.03]	-18.26	[-19.69, -16.83]
7	2022-2019	Not Eligible	-19.86	[-20.31, -19.42]	-18.66	[-19.10, -18.22]
8	2022-2019	Free	-19.24	[-19.64, -18.84]	-14.75	[-15.11, -14.39]
8	2022-2019	Reduced	-20.68	[-22.24, -19.13]	-15.54	[-16.96, -14.12]
8	2022-2019	Not Eligible	-21.25	[-21.71, -20.78]	-15.28	[-15.71, -14.86]

Appendix B

Group Comparisons on the National Assessment of Educational Progress (NAEP) 2022 Mathematics and Reading Grades 4 and 8 assessments

We present student group comparisons on NAEP 2022 mathematics and reading grades 4 and 8 assessments below.

Comparisons by gender

In 2019, female students performed equally well as their male counterparts on NAEP mathematics grade 4 assessments (2019 scale score difference of zero for mathematics grade 4 in Figure 15). The 2022 mean scale score for female and male students for this grade declined by seven and three points, respectively. While 2022 scale scores declined for both groups, the decline in the scale score for female students was larger than the decline in the scale score for male students. This resulted in a net change of four scale score points lower for female students compared to male students in 2022 (mathematics grade 4 panel in Figure 15).

For NAEP reading grade 4, female students' 2019 mean scale score was six points larger than male students' 2019 scale scores. Both gender groups experienced declines in their 2022 NAEP reading scale scores, more so for female students than for male students (6-point decline for female students, 3-point decline for female students). Scale score differences changed from six points in 2019 to three points in 2022, resulting in a net change in scale scores of three points between the two groups in 2022 (reading grade 4 panel in Figure 15). So, female students experienced larger declines in their reading grade 4 performances compared to male students.

In 2019, female students' math grade 8 mean scale score was four points larger than male students' scale scores. Female and male students experienced 11-point and 7-point declines in their 2022 scale scores, respectively, resulting in a change in scale score differences between groups from four to zero points from 2019 to 2022 (mathematics grade 8 panel in Figure 15). Again, female students had larger declines in their math grade 8 performances compared to male students.

In 2019, female students' reading grade 8 mean scale score was 10 points larger than male students' mean scale score. 2022 mean scale scores for female and male students declined by four points and five points, respectively, resulting in a change in scale score differences between groups from 10 to 11 points from 2019 to 2022 (reading grade 8 panel in Figure 15). Male students experienced slightly larger declines in their reading grade 8 performances compared to female students.
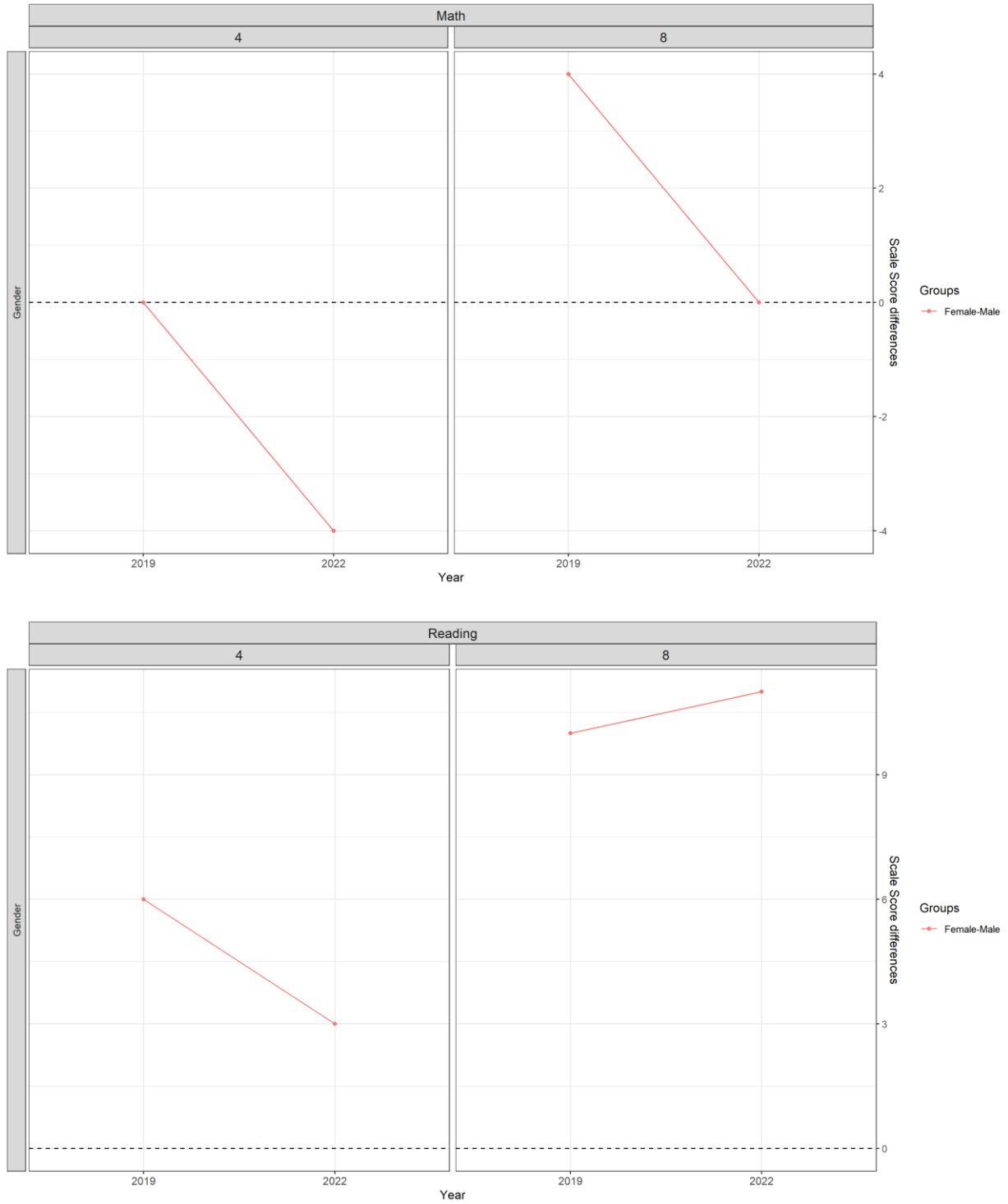
Figure 15. Results by gender from NAEP 2022 mathematics and reading grades 4 and 8 assessments

Comparisons by ethnicity

For NAEP mathematics grade 4, the 2019 mean scale score for Black students was 19 points smaller than the score for White students. The 2019 mean scale score for Hispanic students were 14 points smaller than the score for White students. The 2019 mean scale score for Black students were five points smaller than mean scale scores for Hispanic students. Black, Hispanic, and White groups' 2022 scale scores declined by 13, one, and five points. These scale score changes resulted in 2022 score differences of 27 points smaller for Black students when compared to White students, 10 points smaller for Hispanic students when compared to White students, and 17 points smaller for Black students when compared to Hispanic students (mathematics grade 4 panel in Figure 16). To compare, the NAEP 2019 math grade 4 mean scale score for Black students was smaller than the mean scale score for either White or Hispanic students. Larger 2022 scale score differences for these group comparisons indicated that performance differences between Black and White students and between Black and Hispanic students widened, with Black students performing worse in both comparisons. The 2022 scale score difference between Hispanic and White students was smaller than the 2019 scale score difference by four points, indicating that the performance difference between these groups narrowed in 2022.

The 2019 reading grade 4 mean scale score for Black students was 26 points smaller than the score for White students. The 2019 mean scale score for Hispanic students was 19 points smaller than the mean score for White students. The 2019 mean scale score for Black students was seven points smaller than the scale score for Hispanic students. Black and Hispanic students' 2022 mean scale scores increased by one point, respectively. However, White students' 2022 mean scale score declined by five points. These scale score changes resulted in 2022 score differences of 20 points smaller for Black students when compared to White students, 13 points smaller for Hispanic students when compared to White students, and seven points smaller for Black students when compared to Hispanic students. To compare, the NAEP 2019 reading grade 4 mean scale score for Black students was smaller than scale scores for either White or Hispanic students. Scale score differences between Black and White students and between Hispanic and White students were smaller in 2022 compared to 2019, indicating that performance differences between these two groups decreased in 2022. The 2019 mean scale score difference between Black and Hispanic students remain unchanged in 2022 (indicated by a horizontal line for the comparison between scales scores for Black and Hispanic students in the reading grade 4 panel in Figure 16).

For mathematics grade 8, the 2019 mean scale score for Black students was 26 points smaller than the mean score for White students, and the mean scale score for Hispanic students was 12 points smaller than the mean scale score for White students. Additionally, the mean scale score for Black students was 14 points smaller than the mean scale score for Hispanic students. 2022 scale scores for Black, Hispanic, and White students declined by eight, 11, and eight points, respectively. These scale score changes resulted in 2022 score differences of 26 points smaller for Black students when compared to White students, 15 points smaller for Hispanic when compared to White students, and 11 points smaller for Black students when compared to Hispanic students (mathematics grade 8 panel in Figure 16). The NAEP 2019 math grade 8 mean scale score for Black students was smaller than the mean scale score for either White or Hispanic students. The 2019 scale score difference between Black and White students remain unchanged in 2022 (indicated by a horizontal line for the comparison between scales scores for Black and White students in the mathematics grade 8 panel in Figure 16). The 2022 mean scale score difference between Hispanic and White students was larger than the 2019 mean scale score difference for the same comparison groups, indicating that performance differences

between these two groups increased in 2022. On the other hand, the 2022 scale score difference between Black and Hispanic groups was smaller than the 2019 scale score difference, indicating that performance differences between these two groups decreased in 2022.

The 2019 reading grade 8 mean scale score for Black students was 27 points smaller than the mean score for White students. The 2019 mean scale score for Hispanic students was eight points smaller than the score for White students. The 2019 mean scale score for Black students was 11 points smaller than the score for Hispanic students. Black students' 2022 mean scale score remain unchanged from their 2019 mean scale score while White and Hispanic students' 2022 scale scores declined by four points, respectively. These scale score changes resulted in 2022 score differences of 23 points lower for Black students when compared to White students, eight points lower for Hispanic students when compared to White students, and 15 points lower for Black students when compared to Hispanic students (reading grade 8 panel in Figure 16). The NAEP 2019 reading grade 8 mean scale score for Black students was smaller than the scale score for either White or Hispanic students. 2022 scale score differences between Black and White students and between Black and Hispanic students were smaller than their respective 2019 scale score differences, indicating that performance differences between these two groups decreased in 2022. The 2019 scale score difference between Hispanic and White students remain unchanged in 2022 (indicated by a horizontal line for the comparison between scales scores for Hispanic and White students in the reading grade 8 panel in Figure 16).
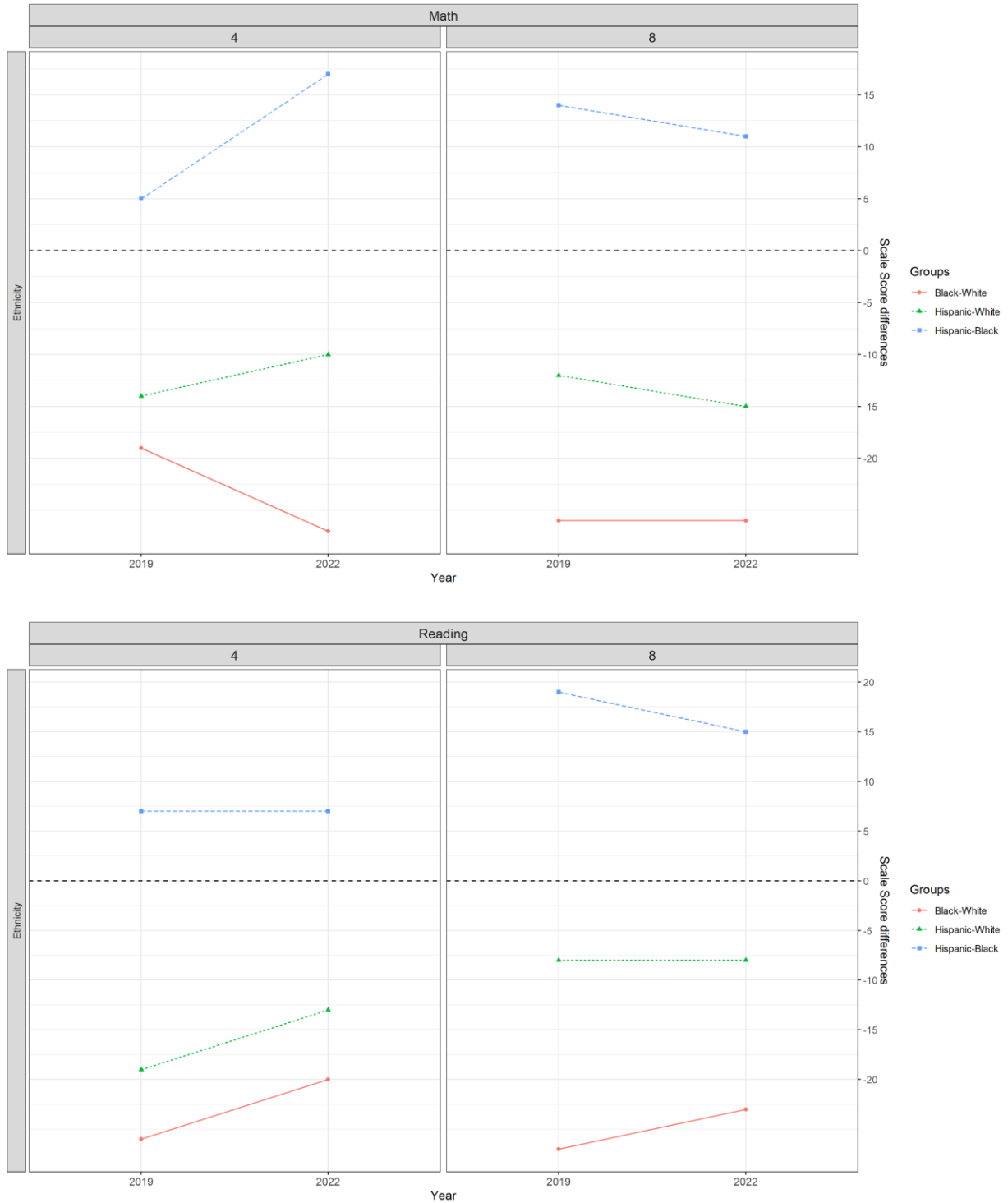
Figure 16. Results by ethnicity from NAEP 2022 mathematics and reading grades 4 and 8 assessments

Comparisons by eligibility for free lunch

The 2019 mathematics grade 4 mean scale score for students who were eligible for free lunch (the *Lunch eligible* group in Table 5) was 23 points smaller than the score for students who were not eligible (the *Not lunch eligible* group in Table 5). The 2022 mean scale score for students in the *Lunch eligible* and the *Not lunch eligible* groups declined by 6 points and 5 points, respectively. These scale score changes resulted in 2022 score differences of 24 points lower for the *Lunch eligible* group when compared to the *Not lunch eligible* group (mathematics grade 4 panel in Figure 17). The NAEP 2019 mathematics grade 4 mean scale score for students eligible for free lunch was smaller than the scale score for students who were not eligible. 2022 scale score differences between the two groups increased by one point, indicating that performance differences between the two groups remained generally stable from 2019 to 2022.

The 2019 reading grade 4 mean scale score for students who were eligible for free lunch was 23 points smaller than the mean score for students who were not eligible. The 2022 scale score for students in the *Lunch eligible* and the *Not lunch eligible* groups declined by 6 points and 3 points, respectively. These mean scale score changes resulted in 2022 score differences of 26 points lower for the *Lunch eligible* group when compared to the *Not lunch eligible* group (reading grade 4 panel in Figure 17). The NAEP 2019 reading grade 8 mean scale score for students eligible for free lunch was smaller than the scale score for students who were not eligible. 2022 scale score differences between the two groups increased by three points, indicating that 2022 performance differences between the two groups remained generally unchanged from 2019.

In 2019, math grade 8 mean scale score for students in the *Lunch eligible* group was 25 points smaller than the scale score for students in the *Not lunch eligible* group. The 2022 scale scores for students in the *Lunch eligible* and *Not lunch eligible* groups declined by eight and nine points, respectively, resulting in 2022 scale score differences of 24 points (mathematics grade 8 panel in Figure 17). The NAEP 2019 mathematics grade 8 mean scale score for students eligible for free lunch was smaller than the scale score for students who were not eligible. 2022 scale score differences between the two groups decreased by one point, indicating that performance differences between the two groups remained generally stable from 2019 to 2022.

In 2019, reading grade 8 mean scale score for students in the *Lunch eligible* group was 24 points smaller than the scale score for students in the *Not lunch eligible* group. 2022 scale scores for the *Lunch eligible* and the *Not lunch eligible* groups declined by four and five points, respectively, resulting in 2022 scale score differences of 23 points (reading grade 8 panel in Figure 17). The NAEP 2019 reading grade 8 mean scale score for students eligible for free lunch was smaller than the scale score for students who were not eligible. 2022 scale score differences between the two groups decreased by one point, indicating that performance differences between the two groups remained generally unchanged from 2019 to 2022.
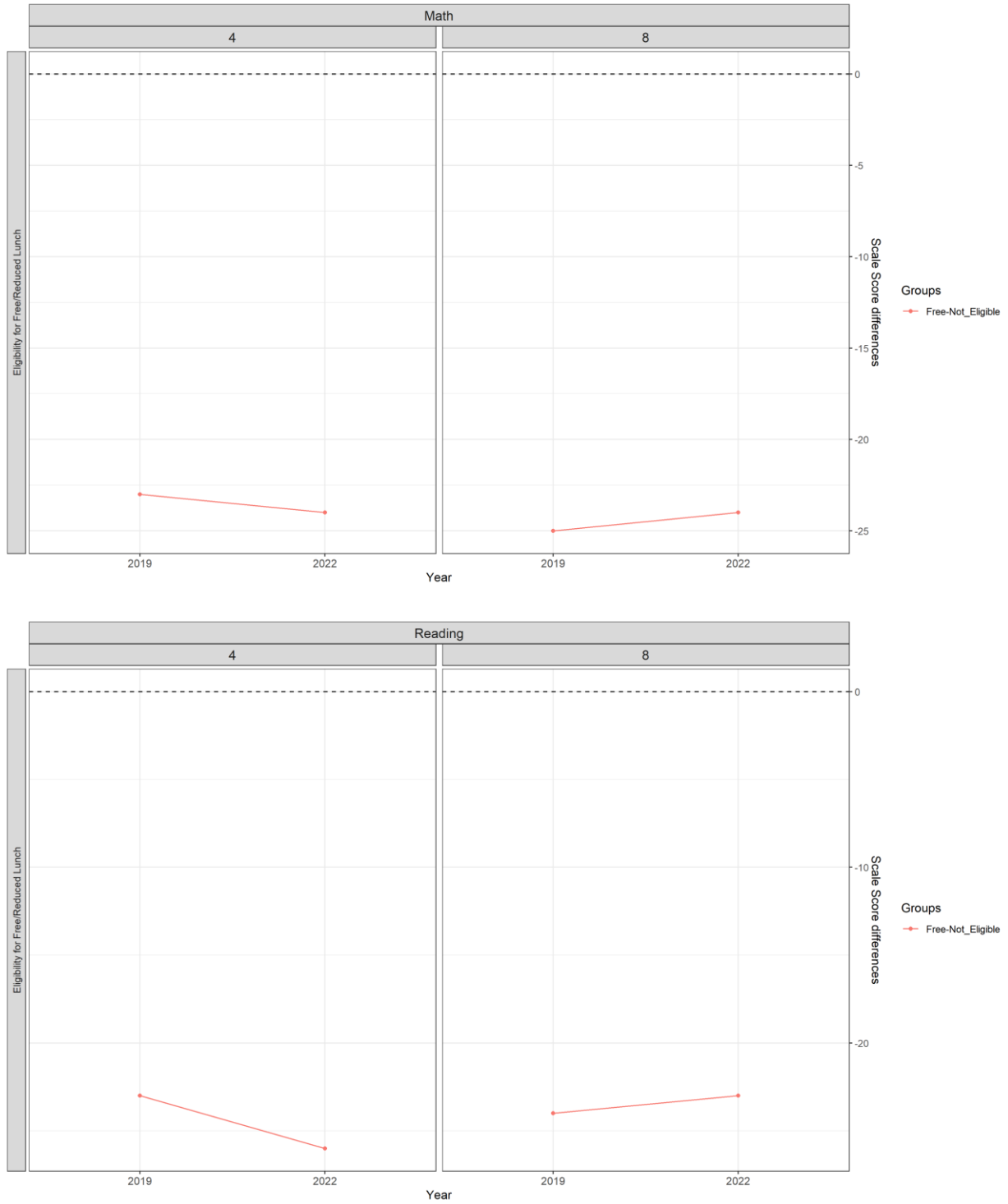
Figure 17. Results by eligibility for free lunch from NAEP 2022 mathematics and reading grades 4 and 8 assessments