California Department of Education (CDE) logo. It is round with a few books and an oil light in the middle.



# Analysis of Remote Testing for the Optional Fall 2019–2020 Summative English Language Proficiency Assessments for California

**Contract #CN140284**

**Prepared for the California Department of Education by ETS**

**Updated March 12, 2021**



Table of Contents

[Section 1: Executive Summary 1](#_Toc66443730)

[1.1. Overview 1](#_Toc66443731)

[1.2. Analysis Sample 1](#_Toc66443732)

[1.3. Student Performance on Individual Test Questions 2](#_Toc66443733)

[1.4. Student Performance on the Entire Test 3](#_Toc66443734)

[1.5. Summary of Validity and Reliability Analyses 3](#_Toc66443735)

[Section 2: Introduction and Purpose 4](#_Toc66443736)

[2.1. Purpose of the Optional Fall Administration 4](#_Toc66443737)

[2.2. Remote Assessment 4](#_Toc66443738)

[2.3. Overview of the Report 4](#_Toc66443739)

[Section 3: Sample Selection 5](#_Toc66443740)

[3.1. Sample Weighting 12](#_Toc66443741)

[3.2. Target Population 13](#_Toc66443742)

[3.3. Weighted Fall 2020 Samples 17](#_Toc66443743)

[Reference 20](#_Toc66443744)

[Section 4: Score Comparability Results 21](#_Toc66443745)

[4.1. Item Performance Comparisons 21](#_Toc66443746)

[4.2. Differential Item Functioning 21](#_Toc66443747)

[4.3. Comparison of Score Means 43](#_Toc66443748)

[References 45](#_Toc66443749)

[Section 5: Supplemental Analysis Results 46](#_Toc66443750)

[5.1. Response Time 46](#_Toc66443751)

[5.2. Score Reliability 53](#_Toc66443752)

[5.3. Correlations Among Domain Scores 55](#_Toc66443753)

[Section 6: Conclusion 59](#_Toc66443754)

[Section 7: Appendix 60](#_Toc66443755)

List of Tables

[Table 3.1 Numbers of Students in the Analysis Sample 5](#_Toc66443756)

[Table 3.2 Percentages of Fall 2020 Students by Group Before Weighting: Oral Language 6](#_Toc66443757)

[Table 3.3 Percentages of Fall 2020 Students by Group Before Weighting: Written Language 9](#_Toc66443758)

[Table 3.4 Descriptive Statistics for 2018–2019 Students by Group: Oral Language 13](#_Toc66443759)

[Table 3.5 Descriptive Statistics for 2018–2019 Students by Group: Written Language 15](#_Toc66443760)

[Table 3.6 Mean ELPAC Scores After Weighting: Oral Language 17](#_Toc66443761)

[Table 3.7 Mean ELPAC Scores After Weighting: Written Language 18](#_Toc66443762)

[Table 3.8 Consistency of Demographic Representation After Weighting: Oral Language 18](#_Toc66443763)

[Table 3.9 Consistency of Demographic Representation After Weighting: Written Language 19](#_Toc66443764)

[Table 4.1 DIF Categories for Dichotomous Items 21](#_Toc66443765)

[Table 4.2 DIF Categories for Polytomous Items 22](#_Toc66443766)

[Table 4.3 Differential Item Functioning Results: Oral Language 23](#_Toc66443767)

[Table 4.4 Differential Item Functioning Results: Written Language 23](#_Toc66443768)

[Table 4.5 Item Difficulty Means: Oral Language 25](#_Toc66443769)

[Table 4.6 Item Difficulty Means: Written Language 25](#_Toc66443770)

[Table 4.7 Item-Total Correlation Means: Oral Language 34](#_Toc66443771)

[Table 4.8 Item-Total Correlation Means: Written Language 34](#_Toc66443772)

[Table 4.9 Scale Score Mean Differences: Oral Language Composite 43](#_Toc66443773)

[Table 4.10 Scale Score Mean Differences: Written Language Composite 44](#_Toc66443774)

[Table 5.1 Response-Time Results: Oral Language, Grade Span Three Through Five 47](#_Toc66443775)

[Table 5.2 Response-Time Results: Oral Language, Grade Span Six Through Eight 48](#_Toc66443776)

[Table 5.3 Response-Time Results: Oral Language, Grade Span Nine and Ten 49](#_Toc66443777)

[Table 5.4 Response Time Results: Written Language, Grade Span Three Through Five 50](#_Toc66443778)

[Table 5.5 Response-Time Results: Written Language, Grade Span Six Through Eight 51](#_Toc66443779)

[Table 5.6 Response-Time Results: Written Language, Grade Span Nine and Ten 52](#_Toc66443780)

[Table 5.7 Reliability: Listening 53](#_Toc66443781)

[Table 5.8 Reliability: Speaking 53](#_Toc66443782)

[Table 5.9 Reliability: Reading 54](#_Toc66443783)

[Table 5.10 Reliability: Writing 54](#_Toc66443784)

[Table 5.11 Correlations Among Domain Raw Scores: Oral Language 55](#_Toc66443785)

[Table 5.12 Correlations Among Domain Raw Scores: Written Language 57](#_Toc66443786)

[Table A.1 Oral Language Classical Item Statistics 61](#_Toc66443787)

[Table A.2 Written Language Classical Item Statistics 71](#_Toc66443788)

[Table A.3 Written Language Kindergarten In Person Raw Scores 80](#_Toc66443789)

[Table A.4 Written Language Kindergarten Remote Raw Scores 83](#_Toc66443790)

List of Figures

[Figure 4.1 Item Difficulty Results: Oral Language, Kindergarten (N=797/819) 26](#_Toc66443791)

[Figure 4.2 Item Difficulty Results: Oral Language, Grade One (N=2,035/1,201) 27](#_Toc66443792)

[Figure 4.3 Item Difficulty Results: Oral Language, Grade Two (N=2,265/1,522) 27](#_Toc66443793)

[Figure 4.4 Item Difficulty Results: Oral Language, Grade Span Three Through Five (N=5,492/3,329) 28](#_Toc66443794)

[Figure 4.5 Item Difficulty Results: Oral Language, Grade Span Six Through Eight (N=3,077/1,583) 28](#_Toc66443795)

[Figure 4.6 Item Difficulty Results: Oral Language, Grade Span Nine and Ten (N=1,317/514) 29](#_Toc66443796)

[Figure 4.7 Item Difficulty Results: Oral Language, Grade Eleven (N=380/154) 29](#_Toc66443797)

[Figure 4.8 Item Difficulty Results: Written Language, Kindergarten (N=131/493) 30](#_Toc66443798)

[Figure 4.9 Item Difficulty Results: Written Language, Grade One (N=198/829) 30](#_Toc66443799)

[Figure 4.10 Item Difficulty Results: Written Language, Grade Two (N=241/905) 31](#_Toc66443800)

[Figure 4.11 Item Difficulty Results: Written Language, Grade Span Three Through Five (N=6,726/4,240) 31](#_Toc66443801)

[Figure 4.12 Item Difficulty Results: Written Language, Grade Span Six Through Eight (N=3,336/2,294) 32](#_Toc66443802)

[Figure 4.13 Item Difficulty Results: Written Language, Grade Span Nine and Ten (N=1,383/724) 32](#_Toc66443803)

[Figure 4.14 Item Difficulty Results: Written Language, Grade Eleven (N=430/217) 33](#_Toc66443804)

[Figure 4.15 Item-Total Correlation Results: Oral Language, Kindergarten (N=797/819) 35](#_Toc66443805)

[Figure 4.16 Item-Total Correlation Results: Oral Language, Grade One (N=2,035/1,201) 36](#_Toc66443806)

[Figure 4.17 Item-Total Correlation Results: Oral Language, Grade Two (N=2,265/1,522) 36](#_Toc66443807)

[Figure 4.18 Item-Total Correlation Results: Oral Language, Grade Span Three Through Five (N=5,492/3,329) 37](#_Toc66443808)

[Figure 4.19 Item-Total Correlation Results: Oral Language, Grade Span Six Through Eight (N=3,077/1,583) 37](#_Toc66443809)

[Figure 4.20 Item-Total Correlation Results: Oral Language, Grade Span Nine and Ten (N=1,317/514) 38](#_Toc66443810)

[Figure 4.21 Item-Total Correlation Results: Oral Language, Grade Eleven (N=380/154) 38](#_Toc66443811)

[Figure 4.22 Item-Total Correlation Results: Written Language, Kindergarten (N=131/493) 39](#_Toc66443812)

[Figure 4.23 Item-Total Correlation Results: Written Language, Grade One (N=198/829) 39](#_Toc66443813)

[Figure 4.24 Item-Total Correlation Results: Written Language, Grade Two (N=241/905) 40](#_Toc66443814)

[Figure 4.25 Item-Total Correlation Results: Written Language, Grade Span Three Through Five (N=6,726/4,240) 40](#_Toc66443815)

[Figure 4.26 Item-Total Correlation Results: Written Language, Grade Span Six Through Eight (N=3,336/2,294) 41](#_Toc66443816)

[Figure 4.27 Item-Total Correlation Results: Written Language, Grade Span Nine and Ten (N=1,383/724) 41](#_Toc66443817)

[Figure 4.28 Item-Total Correlation Results: Written Language, Grade Eleven (N=430/217) 42](#_Toc66443818)

[Figure 5.1 Response-Time Results: Oral Language, Grade Span Three Through Five 47](#_Toc66443819)

[Figure 5.2 Response-Time Results: Oral Language, Grade Span Six Through Eight 48](#_Toc66443820)

[Figure 5.3 Response-Time Results: Oral Language, Grade Span Nine and Ten 49](#_Toc66443821)

[Figure 5.4 Response-Time Results: Written Language, Grade Span Three Through Five 50](#_Toc66443822)

[Figure 5.5 Response-Time Results: Written Language, Grade Span Six Through Eight 51](#_Toc66443823)

[Figure 5.6 Response-Time Results: Written Language, Grade Span Nine and Ten 52](#_Toc66443824)

[Figure 5.7 Follow-up for Correlations: Oral Language, Grade Two 2018–2019 55](#_Toc66443825)

[Figure 5.8 Follow-up for Correlations: Oral Language, Grade Two In Person 56](#_Toc66443826)

[Figure 5.9 Follow-up for Correlations: Oral Language, Grade Two Remote 56](#_Toc66443827)

[Figure 5.10 Follow-up for Correlations: Written Language, Kindergarten In Person 57](#_Toc66443828)

[Figure 5.11 Follow-up for Correlations: Written Language, Kindergarten Remote 58](#_Toc66443829)

## Executive Summary

### Overview

The annual Summative English Language Proficiency Assessments for California (ELPAC) administration window opens on February 1 and ends on May 31. However, in March of 2020, school closures associated with the novel coronavirus disease 2019 (COVID-19) resulted in incomplete testing for many students who were required to be assessed with the Summative ELPAC. With the start of the new school year in fall 2020, the spring 2020 Summative ELPAC administration was extended to August 20 through October 30, 2020.

While the primary purpose of the optional fall administration was to support reclassification, testing was not limited to those students. Because students were not returning to school in a uniform way in the fall of 2020, flexible testing options were allowed. This flexibility meant that students could be tested in two different ways:

1. Students could test in person, with both students and test examiners co-located and following physical distancing guidelines.
2. Students could test remotely, using distance learning tools such as Zoom.

In both scenarios, testing was standardized; for example, test examiners received scripts, called *Directions for Administration (DFAs),* to ensure that students were given the same clear, accurate instructions. The *DFAs* for the Summative ELPAC are prescribed to ensure that students test under conditions that are standardized and as similar as possible. The purpose of this study is to evaluate whether the performance of students testing remotely could be viewed as comparable to the performance of students testing in person.

The ELPAC consists of a written language composite (Reading and Writing domains) and an oral language composite (Speaking and Listening domains). The evaluation includes comparisons of the performance of students, testing in person and remotely, at the composite level and item level; as well as the response time, test reliability, and correlations between domains. The results of the evaluation support the conclusion that remote testing can be viewed as reasonably comparable to in-person testing for the Summative ELPAC under the unusual conditions and with the atypical types of student samples observed here. Further evidence should be gathered before more general and definitive judgments about comparability for the Summative ELPAC or other test populations are made.

### Analysis Sample

Results from more than 20,000 students across grade levels were used in this evaluation. Only students who completed a Summative ELPAC composite during the optional fall administration, tested entirely in one location (either in person or remotely), and were in kindergarten through grade eleven in spring 2020 were included in the analysis sample. This resulted in an oral language performance sample for each grade level that ranged between 380 and 2,265 students who tested remotely across grade levels and between 154 and 1,522 students who tested in person across grade levels. This resulted in a written language performance sample for each grade level that ranged between 131 and 2,689 students who tested remotely across grade levels and between 217 and 1,558 students who tested in person across grade levels.

These numbers represented no more than 3 percent of students who typically participate in ELPAC administrations across grade levels. Very few students who were in grade twelve during the spring administration participated in the optional fall administration, presumably because they graduated. Therefore, grade twelve students were not included in the study.

To facilitate direct comparisons between students who tested in person and students who tested remotely, students in these groups for grades one through eleven were matched based on demographic characteristics and 2018–2019 ELPAC scores to ensure the groups were comparable. Note that very few kindergarten students had 2018–2019 scores; therefore, it was not possible to include the previous year’s ELPAC scores in the matching of kindergarten student groups.

Because prior-year test scores are strong predictors of current-year test scores, the expectation was that matching would produce consistent average performance in grades one through eleven even if remote testing did not impact scores. The expectation was that the matching of in-person and remote groups would be more limited in effectiveness for kindergarten students because demographic variables alone are weaker predictors of test scores.

### Student Performance on Individual Test Questions

The performance of individual test questions, or items, was compared for in-person and remote test takers. Three item analyses were conducted: differential item functioning (DIF), item difficulty, and item discrimination.

DIF methodology was used to identify items that perform differently for in-person and remote test takers. Of the 419 items that were evaluated using DIF across all the tests, only three items were identified as exhibiting the most practically significant form of DIF between the in-person and remote test-taker groups. Each of these items came from a different domain: one in Listening, one in Speaking, and one in Reading. Students testing in person performed better for two of these items, and remote students performed better for the third item.

In addition to DIF, comparisons of item difficulties were conducted. Comparisons of items in oral language for grade one and grade spans three through five, six through eight, and nine and ten were nearly equivalent. Comparisons of item difficulties for written language for grade two, grade span three through five, grade span six through eight, grade span nine and ten, and grade eleven were also nearly equivalent. For both kindergarten oral language and kindergarten written language composites, item difficulty statistics suggested that items were uniformly easier for remote test takers than for in-person test takers. These results may be related to the inability to match based on previous test scores, leading to groups that are less comparable. For grade two oral language, item difficulty statistics suggested that Listening items were more difficult for remote test takers than for in-person test takers. This difference summed to approximately one-half point of the 50 points that could be earned on the test.

Item discrimination is a measure of how strongly item performance is related to test performance. It answers the question, “Are students who correctly answer a question more likely to earn higher scores on the test as a whole than students who incorrectly answer that question?” Differences in item discrimination can mean that the skill being evaluated is changing, so it is important for item discrimination to remain consistent for in-person and remote test takers. Item discrimination results were compared, and results were fairly consistent for in-person and remote test takers, with differences appearing when the number of test takers in a grade level or grade span was small. The degree of consistency increased as the number of students who tested in a grade level or grade span increased.

### Student Performance on the Entire Test

Mean Summative ELPAC scores for students in the in-person and remote groups were compared using statistical tests to evaluate whether differences were larger than differences that might be expected by chance alone. In the study, both oral language composite scores and written language composite scores in each of the seven grade levels and grade spans—kindergarten, grade one, grade two, grade span three through five, grade span six through eight, grade span nine and ten, and grade eleven—were evaluated, for a total of 14 composite scores.

On the basis of these comparisons, no differences were found for 12 of the 14 ELPAC scores. Kindergarten oral language scores were 17.1 scale score points higher for remote test takers than for in-person test takers. The practical significance of this difference is small, and this result is consistent with the earlier finding that items were less difficult for kindergarten remote test takers than for kindergarten in-person test takers. As in that case, matching that was limited to demographic group membership is a potential explanation for this result. Grade two oral language scores were 8.6 scale score points lower for remote test takers than for in-person test takers; the practical significance of this difference is also small. The remaining 12 analyses showed no differences across groups.

### Summary of Validity and Reliability Analyses

The validity and reliability analyses suggested no threats to the psychometric properties of the remotely administered tests. Response times were a bit longer for remote test takers, but nothing in these differences suggested anything problematic in the test delivery scenarios. Reliability estimates were similar for in-person and remote test-taker scores; differences between them were associated with differences in score variability, suggesting that reduced variability was responsible for the differences. Differences in correlations between domains were present but small and followed no pattern that might have suggested a problem in what the Summative ELPAC measures across the two modes of test administration.

## Introduction and Purpose

This chapter provides an overview of the optional fall administration of the 2019–2020 Summative ELPAC, including the purpose of the administration, the testing window, and an overview of the report.

### Purpose of the Optional Fall Administration

Because of COVID-19, the spring 2020 Summative ELPAC administration was suspended in March 2020. An optional fall administration was offered between August 20 and October 30, 2020, for students who were unable to start or complete the test in spring 2020. This administration used the same test forms as were used in the spring. These students tested off-grade in that grade three students, for example, were administered the grade two assessment that would have been administered in spring 2020.

### Remote Assessment

Schools could test in two possible locations during the optional fall administration:

* Testing could be in person, co-located with test examiners but conforming to physical distancing guidelines
* Testing could be remote using distance learning tools such as Zoom or Google Classroom

The aim of standardized testing is to minimize opportunities for differences in test administration to impact scores. Allowing for remote testing introduced a challenge to the consistency of scores that were obtained from the fall 2020 administrations. The purpose of this study was to evaluate the consistency of in-person and remote testing outcomes.

### Overview of the Report

This report contains information about the study participants, a comparison of test performance information for in-person and remote test takers, and some supplemental information that supports the test performance information.

Study participants are described in terms of demographic composition and scale scores. A matching approach is described, and its results are provided.

Test performance is compared in terms of item outcomes and test score outcomes.

Supplemental information provides other perspectives on student outcomes to inform whether in-person and remote scores can be viewed as being comparable.

## Sample Selection

This chapter describes the students who participated in the optional fall 2020 Summative ELPAC administration and who are included in the study, and the approach used to weight these students’ scores to allow for remote testing to be evaluated.

The ELPAC is composed of two parts: oral language and written language. The oral language composite consists of the Listening and Speaking domains. The written language composite consists of the Reading and Writing domains. Table 3.1 contains the number of students who participated in the oral language and written language portions of the optional fall 2020 Summative ELPAC administration. To be included, these students needed to have started and completed testing in fall 2020 for both domains in the composite.

Within the table, values associated with fewer than 500 students being in a grade level or grade span are bolded and have an asterisk. This is because variations in test outcomes are expected when fewer than 500 students participate in an administration. Overall, few fall 2020 kindergarten students tested in written language composite. For the remote testing location, the number of students tested in kindergarten through grade two written language was much smaller than the other grade levels or grade spans. This may be because the written test is a paper–pencil test for kindergarten through grade two, and schools were uncomfortable with delivering secure paper–pencil Writing materials to students.

Fewer than 500 students in grade eleven participated in both oral language and written language parts in either the in-person or remote locations. Because few students in grade twelve would be returning to be tested in the fall, those students were not evaluated.

**Table 3.1 Numbers of Students in the Analysis Sample**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Grade Level | Oral Language In Person | Oral Language Remote | Written Language In Person | Written Language Remote |
| Kindergarten | 819 | 797 | **493** | **131** |
| 1 | 1,201 | 2,035 | 829 | **198** |
| 2 | 1,522 | 2,265 | 905 | **241** |
| 3 | 1,285 | 2,259 | 1,558 | 2,689 |
| 4 | 959 | 1,828 | 1,225 | 2,191 |
| 5 | 1,085 | 1,405 | 1,457 | 1,846 |
| 6 | 592 | 1,289 | 973 | 1,439 |
| 7 | 582 | 1,019 | 792 | 1,081 |
| 8 | 409 | 769 | 529 | 816 |
| 9 | 310 | 760 | 428 | 803 |
| 10 | 204 | 557 | 296 | 580 |
| 11 | **\*154** | **\*380** | **\*217** | **\*430** |

Table 3.2 includes the numbers of all oral language students in each grade and for groups of students within those grades for both in-person and remote conditions. It also contains the differences in the percentage representation and differences in the scale score means before weighting was conducted. Note that differences are present, but those differences are often quite small when sample sizes are reasonable (e.g., 500).

Table 3.2 Percentages of Fall 2020 Students by Group Before Weighting: Oral Language

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | Group | In Person N | Remote N | Percent in Group Difference | In Person Mean Scale Score | Remote Mean Scale Score | Mean Scale Score Difference |
| Kindergarten | All | 819 | 797 | N/A | 1474 | 1496 | 22 |
| Kindergarten | Male | 422 | 381 | -4 | 1469 | 1489 | 20 |
| Kindergarten | Female | 397 | 416 | 4 | 1479 | 1502 | 23 |
| Kindergarten | Asian | 108 | 52 | -6 | 1496 | 1535 | 39 |
| Kindergarten | Hispanic or Latino | 657 | 693 | 7 | 1469 | 1491 | 22 |
| Kindergarten | Special education services | 54 | 71 | 2 | 1444 | 1458 | 14 |
| Kindergarten | Economically disadvantaged | 476 | 589 | 16 | 1466 | 1493 | 27 |
| Kindergarten | Parent/Guardian education: some college or higher | 553 | 213 | -41 | 1481 | 1487 | 6 |
| 1 | All | 1,201 | 2,035 | N/A | 1492 | 1492 | 0 |
| 1 | Male | 612 | 1,072 | 2 | 1492 | 1491 | -1 |
| 1 | Female | 589 | 963 | -2 | 1492 | 1492 | 0 |
| 1 | Asian | 174 | 135 | -7 | 1504 | 1518 | 14 |
| 1 | Hispanic or Latino | 950 | 1,792 | 9 | 1488 | 1489 | 1 |
| 1 | Special education services | 69 | 191 | 3 | 1469 | 1481 | 12 |
| 1 | Economically disadvantaged | 654 | 1,660 | 28 | 1485 | 1489 | 4 |
| 1 | Parent/Guardian education: some college or higher | 780 | 391 | -46 | 1496 | 1496 | 0 |
| 2 | All | 1,522 | 2,265 | N/A | 1518 | 1514 | -4 |
| 2 | Male | 786 | 1,158 | -1 | 1519 | 1516 | -3 |
| 2 | Female | 736 | 1,107 | 1 | 1516 | 1513 | -3 |
| 2 | Asian | 205 | 186 | -5 | 1540 | 1540 | 0 |
| 2 | Hispanic or Latino | 1,243 | 1,939 | 4 | 1513 | 1510 | -3 |
| 2 | Special education services | 63 | 167 | 3 | 1486 | 1504 | 18 |
| 2 | Economically disadvantaged | 899 | 1,727 | 17 | 1513 | 1511 | -2 |
| 2 | Parent/Guardian education: some college or higher | 939 | 614 | -35 | 1522 | 1520 | -2 |

Table 3.2 *(continuation one)*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | Group | In Person N | Remote N | Percent in Group Difference | In Person Mean Scale Score | Remote Mean Scale Score | Mean Scale Score Difference |
| 3–5 | All | 3,329 | 5,492 | N/A | 1532 | 1531 | -1 |
| 3–5 | Male | 1,687 | 2,790 | 0 | 1536 | 1530 | -6 |
| 3–5 | Female | 1,642 | 2,702 | 0 | 1527 | 1531 | 4 |
| 3–5 | Asian | 208 | 276 | -1 | 1534 | 1556 | 22 |
| 3–5 | Hispanic or Latino | 2,861 | 4,852 | 2 | 1529 | 1528 | -1 |
| 3–5 | Special education services | 293 | 682 | 3 | 1519 | 1518 | -1 |
| 3–5 | Economically disadvantaged | 2,165 | 4,558 | 18 | 1528 | 1528 | 0 |
| 3–5 | Parent/Guardian education: some college or higher | 1,738 | 964 | -34 | 1534 | 1541 | 7 |
| 6–8 | All | 1,583 | 3,077 | N/A | 1573 | 1574 | 1 |
| 6–8 | Male | 835 | 1,565 | -2 | 1574 | 1573 | -1 |
| 6–8 | Female | 748 | 1,512 | 2 | 1573 | 1575 | 2 |
| 6–8 | Asian | 111 | 161 | -2 | 1616 | 1602 | -14 |
| 6–8 | Hispanic or Latino | 1,372 | 2,670 | 0 | 1568 | 1569 | 1 |
| 6–8 | Special education services | 231 | 525 | 2 | 1555 | 1554 | -1 |
| 6–8 | Economically disadvantaged | 1,024 | 2,439 | 14 | 1572 | 1569 | -3 |
| 6–8 | Parent/Guardian education: some college or higher | 783 | 674 | -27 | 1580 | 1585 | 5 |
| 9–10 | All | 514 | 1,317 | N/A | 1576 | 1591 | 15 |
| 9–10 | Male | 291 | 659 | -7 | 1583 | 1595 | 12 |
| 9–10 | Female | 223 | 658 | 7 | 1567 | 1587 | 20 |
| 9–10 | Asian | 37 | 80 | -1 | 1655 | 1650 | -5 |
| 9–10 | Hispanic or Latino | 454 | 1,141 | -1 | 1567 | 1583 | 16 |
| 9–10 | Special education services | 87 | 207 | -1 | 1541 | 1567 | 26 |
| 9–10 | Economically disadvantaged | 334 | 1,004 | 11 | 1566 | 1585 | 19 |
| 9–10 | Parent/Guardian education: some college or higher | 255 | 323 | -25 | 1593 | 1606 | 13 |

Table 3.2 *(continuation two)*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | Group | In Person N | Remote N | Percent in Group Difference | In Person Mean Scale Score | Remote Mean Scale Score | Mean Scale Score Difference |
| 11 | All | 154 | 380 | N/A | 1582 | 1590 | 8 |
| 11 | Male | 79 | 187 | -2 | 1588 | 1596 | 8 |
| 11 | Female | 75 | 193 | 2 | 1576 | 1584 | 8 |
| 11 | Asian | 11 | 23 | -1 | 1683 | 1587 | -96 |
| 11 | Hispanic or Latino | 134 | 316 | -4 | 1577 | 1590 | 13 |
| 11 | Special education services | 27 | 52 | -4 | 1561 | 1560 | -1 |
| 11 | Economically disadvantaged | 89 | 293 | 19 | 1578 | 1587 | 9 |
| 11 | Parent/Guardian education: some college or higher | 76 | 81 | -28 | 1591 | 1590 | -1 |

Table 3.3 provides corresponding information for written language.

Table 3.3 Percentages of Fall 2020 Students by Group Before Weighting: Written Language

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | Group | In Person N | Remote N | Percent in Group Difference | In Person Mean Scale Score | Remote Mean Scale Score | Mean Scale Score Difference |
| Kindergarten | All | 493 | 131 | N/A | 1495 | 1519 | 24 |
| Kindergarten | Male | 256 | 67 | -1 | 1483 | 1516 | 33 |
| Kindergarten | Female | 237 | 64 | 1 | 1508 | 1523 | 15 |
| Kindergarten | Asian | 69 | 25 | 5 | 1548 | 1608 | 60 |
| Kindergarten | Hispanic or Latino | 392 | 93 | -9 | 1484 | 1487 | 3 |
| Kindergarten | Special education services | 28 | 12 | 3 | 1488 | 1529 | 41 |
| Kindergarten | Economically disadvantaged | 300 | 50 | -23 | 1480 | 1512 | 32 |
| Kindergarten | Parent/Guardian education: some college or higher | 329 | 94 | 5 | 1507 | 1530 | 23 |
| 1 | All | 829 | 198 | N/A | 1485 | 1488 | 3 |
| 1 | Male | 412 | 110 | 6 | 1479 | 1486 | 7 |
| 1 | Female | 417 | 88 | -6 | 1490 | 1489 | -1 |
| 1 | Asian | 123 | 23 | -3 | 1520 | 1545 | 25 |
| 1 | Hispanic or Latino | 661 | 164 | 3 | 1476 | 1478 | 2 |
| 1 | Special education services | 47 | 21 | 5 | 1454 | 1470 | 16 |
| 1 | Economically disadvantaged | 444 | 106 | 0 | 1470 | 1477 | 7 |
| 1 | Parent/Guardian education: some college or higher | 556 | 142 | 5 | 1493 | 1489 | -4 |

Table 3.3 *(continuation one)*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | Group | In Person N | Remote N | Percent in Group Difference | In Person Mean Scale Score | Remote Mean Scale Score | Mean Scale Score Difference |
| 2 | All | 905 | 241 | N/A | 1520 | 1528 | 8 |
| 2 | Male | 453 | 139 | 8 | 1517 | 1526 | 9 |
| 2 | Female | 452 | 102 | -8 | 1522 | 1531 | 9 |
| 2 | Asian | 116 | 25 | -3 | 1555 | 1592 | 37 |
| 2 | Hispanic or Latino | 756 | 196 | -3 | 1514 | 1517 | 3 |
| 2 | Special education services | 43 | 18 | 2 | 1490 | 1505 | 15 |
| 2 | Economically disadvantaged | 520 | 107 | -13 | 1513 | 1512 | -1 |
| 2 | Parent/Guardian education: some college or higher | 585 | 171 | 6 | 1524 | 1534 | 10 |
| 3–5 | All | 4,240 | 6,726 | N/A | 1532 | 1524 | -8 |
| 3–5 | Male | 2,163 | 3,381 | -1 | 1528 | 1521 | -7 |
| 3–5 | Female | 2,077 | 3,345 | 1 | 1536 | 1526 | -10 |
| 3–5 | Asian | 320 | 357 | -3 | 1547 | 1557 | 10 |
| 3–5 | Hispanic or Latino | 3,612 | 5,947 | 3 | 1529 | 1520 | -9 |
| 3–5 | Special education services | 376 | 750 | 2 | 1507 | 1510 | 3 |
| 3–5 | Economically disadvantaged | 2,782 | 5,419 | 15 | 1529 | 1520 | -9 |
| 3–5 | Parent/Guardian education: some college or higher | 2,134 | 1,425 | -29 | 1534 | 1541 | 7 |
| 6–8 | All | 2,294 | 3,336 | N/A | 1560 | 1554 | -6 |
| 6–8 | Male | 1,183 | 1,677 | -2 | 1555 | 1550 | -5 |
| 6–8 | Female | 1,111 | 1,659 | 2 | 1564 | 1559 | -5 |
| 6–8 | Asian | 181 | 204 | -2 | 1589 | 1593 | 4 |
| 6–8 | Hispanic or Latino | 1,948 | 2,853 | 1 | 1555 | 1549 | -6 |
| 6–8 | Special education services | 317 | 550 | 2 | 1538 | 1531 | -7 |
| 6–8 | Economically disadvantaged | 1,511 | 2,591 | 12 | 1557 | 1550 | -7 |
| 6–8 | Parent/Guardian education: some college or higher | 1,091 | 838 | -23 | 1564 | 1568 | 4 |

Table 3.3 *(continuation two)*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | Group | In Person N | Remote N | Percent in Group Difference | In Person Mean Scale Score | Remote Mean Scale Score | Mean Scale Score Difference |
| 9–10 | All | 724 | 1,383 | N/A | 1579 | 1582 | 3 |
| 9–10 | Male | 418 | 704 | -7 | 1580 | 1576 | -4 |
| 9–10 | Female | 306 | 679 | 7 | 1578 | 1588 | 10 |
| 9–10 | Asian | 51 | 89 | -1 | 1612 | 1609 | -3 |
| 9–10 | Hispanic or Latino | 630 | 1,186 | -1 | 1575 | 1577 | 2 |
| 9–10 | Special education services | 115 | 211 | -1 | 1555 | 1556 | 1 |
| 9–10 | Economically disadvantaged | 468 | 1,061 | 12 | 1574 | 1579 | 5 |
| 9–10 | Parent/Guardian education: some college or higher | 363 | 334 | -26 | 1587 | 1591 | 4 |
| 11 | All | 217 | 430 | N/A | 1587 | 1590 | 3 |
| 11 | Male | 104 | 214 | 2 | 1579 | 1586 | 7 |
| 11 | Female | 113 | 216 | -2 | 1593 | 1594 | 1 |
| 11 | Asian | 16 | 24 | -1 | 1637 | 1606 | -31 |
| 11 | Hispanic or Latino | 190 | 361 | -4 | 1584 | 1587 | 3 |
| 11 | Special education services | 45 | 54 | -8 | 1563 | 1557 | -6 |
| 11 | Economically disadvantaged | 128 | 332 | 18 | 1582 | 1586 | 4 |
| 11 | Parent/Guardian education: some college or higher | 106 | 91 | -28 | 1591 | 1597 | 6 |

### Sample Weighting

While the student group compositions and scale score means were sometimes similar, at times they were not; this is likely because test takers were not randomly assigned to the in-person or remote groups. Schools or local educational agencies (LEAs) decided whether to test students in the optional fall 2020 administration; those choosing to test decided whether to test the students in person or remotely. This meant that in-person and remote groups could not be assumed to be equivalent in their expected performance.

Minimum discriminant information analysis (MDIA; Haberman, 1984) is an approach to weighting scores that allows percentages of students in student groups and mean scale scores to be aligned with a target sample or population. For the purposes of this study, the 2018–2019 administration was chosen to be the target.

The percentage of 2018–2019 female students, male students, Asian students, Hispanic or Latino students, students receiving special education services, students who are economically disadvantaged, and students whose parents/guardians had at least some college education were targeted. Mean 2018–2019 scale scores were also targeted for grades one through eleven. Scale scores were not targeted for kindergarten because past scores are not available for the great majority of those students.

MDIA was applied to fall 2020 in-person test takers, assigning weights to each student with the aim of matching the percentage of students in each student group and the students’ mean scale scores. MDIA was subsequently applied to fall 2020 remote test takers. The result of this approach was that fall 2020 in-person and fall 2020 remote test takers were both matched with the 2018–2019 population and, therefore, matched with each other.

### Target Population

The 2018–2019 administration was selected as the target population because it was the most recent administration that was completed. That is, the 2019–2020 administration was not selected because only a fraction of students tested, and those students may or may not have been a random sample from the population.

The characteristics of the 2018–2019 population are described in table 3.4 for oral language and table 3.5 for written language.

Table 3.4 Descriptive Statistics for 2018–2019 Students by Group: Oral Language

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | Group | N | Percent in Group | Scale Score Mean | Scale Score SD |
| **Kindergarten** | **All** | **153,500** | **100** | **1435** | **66** |
| Kindergarten | Male | 81,094 | 53 | 1430 | 68 |
| Kindergarten | Female | 72,406 | 47 | 1440 | 64 |
| Kindergarten | Asian | 21,494 | 14 | 1440 | 74 |
| Kindergarten | Hispanic or Latino | 119,044 | 78 | 1433 | 64 |
| Kindergarten | Special education services | 13,748 | 9 | 1385 | 96 |
| Kindergarten | Economically disadvantaged | 120,514 | 79 | 1433 | 63 |
| Kindergarten | Parent/Guardian education: some college or higher | 89,510 | 58 | 1440 | 66 |
| **1** | **All** | **127,401** | **100** | **1460** | **64** |
| 1 | Male | 66,498 | 52 | 1456 | 65 |
| 1 | Female | 60,903 | 48 | 1465 | 61 |
| 1 | Asian | 16,608 | 13 | 1469 | 71 |
| 1 | Hispanic or Latino | 99,928 | 78 | 1458 | 62 |
| 1 | Special education services | 13,353 | 10 | 1413 | 94 |
| 1 | Economically disadvantaged | 104,579 | 82 | 1459 | 60 |
| 1 | Parent/Guardian education: some college or higher | 72,621 | 57 | 1466 | 62 |
| **2** | **All** | **115,637** | **100** | **1486** | **68** |
| 2 | Male | 61,022 | 53 | 1484 | 72 |
| 2 | Female | 54,615 | 47 | 1489 | 64 |
| 2 | Asian | 12,468 | 11 | 1499 | 82 |
| 2 | Hispanic or Latino | 93,654 | 81 | 1484 | 65 |
| 2 | Special education services | 14,572 | 13 | 1437 | 97 |
| 2 | Economically disadvantaged | 97,276 | 84 | 1484 | 64 |
| 2 | Parent/Guardian education: some college or higher | 63,794 | 55 | 1492 | 68 |

Table 3.4 *(continuation)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | Group | N | Percent in Group | Scale Score Mean | Scale Score SD |
| **3–5** | **All** | **291,281** | **100** | **1501** | **76** |
| 3–5 | Male | 157,537 | 54 | 1500 | 79 |
| 3–5 | Female | 133,744 | 46 | 1503 | 71 |
| 3–5 | Asian | 27,083 | 9 | 1506 | 91 |
| 3–5 | Hispanic or Latino | 241,948 | 83 | 1501 | 73 |
| 3–5 | Special education services | 53,200 | 18 | 1463 | 97 |
| 3–5 | Economically disadvantaged | 251,762 | 86 | 1501 | 72 |
| 3–5 | Parent/Guardian education: some college or higher | 150,029 | 52 | 1507 | 75 |
| **6–8** | **All** | **204,780** | **100** | **1522** | **96** |
| 6–8 | Male | 116,651 | 57 | 1522 | 99 |
| 6–8 | Female | 88,129 | 43 | 1522 | 92 |
| 6–8 | Asian | 17,060 | 8 | 1531 | 117 |
| 6–8 | Hispanic or Latino | 173,012 | 84 | 1521 | 93 |
| 6–8 | Special education services | 50,758 | 25 | 1490 | 108 |
| 6–8 | Economically disadvantaged | 178,750 | 87 | 1522 | 92 |
| 6–8 | Parent/Guardian education: some college or higher | 96,676 | 47 | 1531 | 95 |
| **9–10** | **All** | **101,579** | **100** | **1527** | **124** |
| 9–10 | Male | 60,279 | 59 | 1529 | 127 |
| 9–10 | Female | 41,300 | 41 | 1525 | 120 |
| 9–10 | Asian | 9,280 | 9 | 1544 | 139 |
| 9–10 | Hispanic or Latino | 84,456 | 83 | 1524 | 122 |
| 9–10 | Special education services | 24,486 | 24 | 1501 | 131 |
| 9–10 | Economically disadvantaged | 85,771 | 84 | 1527 | 119 |
| 9–10 | Parent/Guardian education: some college or higher | 43,075 | 42 | 1546 | 121 |
| **11** | **All** | **41,813** | **100** | **1520** | **114** |
| 11 | Male | 24,702 | 59 | 1518 | 117 |
| 11 | Female | 17,111 | 41 | 1521 | 110 |
| 11 | Asian | 4,345 | 10 | 1531 | 125 |
| 11 | Hispanic or Latino | 34,213 | 82 | 1517 | 112 |
| 11 | Special education services | 9,942 | 24 | 1487 | 131 |
| 11 | Economically disadvantaged | 35,077 | 84 | 1519 | 111 |
| 11 | Parent/Guardian education: some college or higher | 17,294 | 41 | 1534 | 114 |

Table 3.5 Descriptive Statistics for 2018–2019 Students by Group: Written Language

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | Group | N | Percent in Group | Scale Score Mean | Scale Score SD |
| **Kindergarten** | **All** | **153,500** | **100** | **1405** | **89** |
| Kindergarten | Male | 81,094 | 53 | 1401 | 90 |
| Kindergarten | Female | 72,406 | 47 | 1410 | 87 |
| Kindergarten | Asian | 21,494 | 14 | 1435 | 105 |
| Kindergarten | Hispanic or Latino | 119,044 | 78 | 1399 | 84 |
| Kindergarten | Special education services | 13,748 | 9 | 1355 | 99 |
| Kindergarten | Economically disadvantaged | 120,514 | 79 | 1401 | 84 |
| Kindergarten | Parent/Guardian education: some college or higher | 89,510 | 58 | 1412 | 90 |
| **1** | **All** | **127,401** | **100** | **1446** | **78** |
| 1 | Male | 66,498 | 52 | 1442 | 80 |
| 1 | Female | 60,903 | 48 | 1451 | 75 |
| 1 | Asian | 16,608 | 13 | 1484 | 95 |
| 1 | Hispanic or Latino | 99,928 | 78 | 1438 | 72 |
| 1 | Special education services | 13,353 | 10 | 1391 | 100 |
| 1 | Economically disadvantaged | 104,579 | 82 | 1441 | 72 |
| 1 | Parent/Guardian education: some college or higher | 72,621 | 57 | 1455 | 79 |
| **2** | **All** | **115,637** | **100** | **1479** | **72** |
| 2 | Male | 61,022 | 53 | 1473 | 73 |
| 2 | Female | 54,615 | 47 | 1486 | 71 |
| 2 | Asian | 12,468 | 11 | 1512 | 91 |
| 2 | Hispanic or Latino | 93,654 | 81 | 1473 | 67 |
| 2 | Special education services | 14,572 | 13 | 1424 | 93 |
| 2 | Economically disadvantaged | 97,276 | 84 | 1475 | 68 |
| 2 | Parent/Guardian education: some college or higher | 63,794 | 55 | 1486 | 74 |
| **3–5** | **All** | **291,281** | **100** | **1508** | **66** |
| 3–5 | Male | 157,537 | 54 | 1501 | 68 |
| 3–5 | Female | 133,744 | 46 | 1515 | 62 |
| 3–5 | Asian | 27,083 | 9 | 1519 | 79 |
| 3–5 | Hispanic or Latino | 241,948 | 83 | 1506 | 64 |
| 3–5 | Special education services | 53,200 | 18 | 1463 | 93 |
| 3–5 | Economically disadvantaged | 251,762 | 86 | 1507 | 63 |
| 3–5 | Parent/Guardian education: some college or higher | 150,029 | 52 | 1512 | 65 |

Table 3.5 *(continuation)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | Group | N | Percent in Group | Scale Score Mean | Scale Score SD |
| **6–8** | **All** | **204,780** | **100** | **1528** | **73** |
| 6–8 | Male | 116,651 | 57 | 1522 | 74 |
| 6–8 | Female | 88,129 | 43 | 1534 | 70 |
| 6–8 | Asian | 17,060 | 8 | 1541 | 91 |
| 6–8 | Hispanic or Latino | 173,012 | 84 | 1526 | 70 |
| 6–8 | Special education services | 50,758 | 25 | 1490 | 96 |
| 6–8 | Economically disadvantaged | 178,750 | 87 | 1527 | 69 |
| 6–8 | Parent/Guardian education: some college or higher | 96,676 | 47 | 1534 | 72 |
| **9–10** | **All** | **101,579** | **100** | **1534** | **88** |
| 9–10 | Male | 60,279 | 59 | 1528 | 89 |
| 9–10 | Female | 41,300 | 41 | 1542 | 86 |
| 9–10 | Asian | 9,280 | 9 | 1553 | 99 |
| 9–10 | Hispanic or Latino | 84,456 | 83 | 1531 | 86 |
| 9–10 | Special education services | 24,486 | 24 | 1497 | 114 |
| 9–10 | Economically disadvantaged | 85,771 | 84 | 1534 | 84 |
| 9–10 | Parent/Guardian education: some college or higher | 43,075 | 42 | 1545 | 87 |
| **11** | **All** | **41,813** | **100** | **1542** | **94** |
| 11 | Male | 24,702 | 59 | 1536 | 96 |
| 11 | Female | 17,111 | 41 | 1550 | 89 |
| 11 | Asian | 4,345 | 10 | 1555 | 100 |
| 11 | Hispanic or Latino | 34,213 | 82 | 1540 | 92 |
| 11 | Special education services | 9,942 | 24 | 1498 | 126 |
| 11 | Economically disadvantaged | 35,077 | 84 | 1542 | 90 |
| 11 | Parent/Guardian education: some college or higher | 17,294 | 41 | 1551 | 94 |

### Weighted Fall 2020 Samples

After MDIA weights were assigned, some basic checks were performed to ensure that the weights met their objective of matching fall 2020 test takers to the 2018–2019 population values.

Table 3.6 contains mean oral language raw scores for three groups of students:

* Target—All test takers in 2018–2019 (the target population)
* In Person—Weighted 2020 in-person students
* Remote—Weighted 2020 remote students

The difference column shows that the MDIA procedure was successful in assigning weights that match raw score means. Note that kindergarten is excluded from this table because 2018–2019 scores would not have been available for the large majority of students.

Table 3.6 Mean ELPAC Scores After Weighting: Oral Language

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Grade Level | Target | Weighted Remote | Weighted In Person | Difference (Weighted Remote–Weighted In Person) |
| 1 | 32.44 | 32.44 | 32.44 | 0.00000 |
| 2 | 38.71 | 38.71 | 38.71 | 0.00000 |
| 3 | 36.93 | 36.93 | 36.93 | 0.00000 |
| 4 | 39.91 | 39.91 | 39.91 | 0.00000 |
| 5 | 40.95 | 40.95 | 40.95 | 0.00000 |
| 6 | 38.17 | 38.17 | 38.17 | 0.00000 |
| 7 | 38.70 | 38.70 | 38.70 | 0.00000 |
| 8 | 38.98 | 38.98 | 38.98 | 0.00000 |
| 9 | 35.83 | 35.83 | 35.83 | 0.00000 |
| 10 | 36.85 | 36.85 | 36.85 | 0.00000 |
| 11 | 35.44 | 35.44 | 35.44 | 0.00000 |

Table 3.7 contains corresponding information for written language scores. As with oral language, MDIA was successful in assigning weights that matched students’ 2018–2019 raw scores.

Table 3.7 Mean ELPAC Scores After Weighting: Written Language

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Grade Level | Target | Weighted Remote | Weighted In Person | Difference (Weighted Remote—Weighted In Person) |
| 1 | 23.16 | 23.16 | 23.16 | 0.00000 |
| 2 | 29.41 | 29.41 | 29.41 | 0.00000 |
| 3 | 22.98 | 22.98 | 22.98 | 0.00000 |
| 4 | 27.17 | 27.17 | 27.17 | 0.00000 |
| 5 | 29.22 | 29.22 | 29.22 | 0.00000 |
| 6 | 21.26 | 21.26 | 21.26 | 0.00000 |
| 7 | 22.65 | 22.65 | 22.65 | 0.00000 |
| 8 | 23.77 | 23.77 | 23.77 | 0.00000 |
| 9 | 23.37 | 23.37 | 23.37 | 0.00000 |
| 10 | 24.70 | 24.70 | 24.70 | 0.00000 |
| 11 | 23.95 | 23.95 | 23.95 | 0.00000 |

Both test scores and student group representation were matched using MDIA. Table 3.8 contains the differences between the percentage of students in fall 2020 oral language. Values in the table reflect differences between in-person test takers and the target values. Remote differences are identical. In both cases, differences are smaller than one tenth of one percent.

Table 3.8 Consistency of Demographic Representation After Weighting: Oral Language

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Student Group | Kindergarten | 1 | 2 | 3–5 | 6–8 | 9–10 | 11 |
| All | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Male | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Female | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Asian | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Hispanic or Latino | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Special education services | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Economically disadvantaged | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Parent/Guardian education: some college or higher | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |

Corresponding information is provided in table 3.9 for written language. As with oral language, differences are smaller than one tenth of one percent.

Table 3.9 Consistency of Demographic Representation After Weighting: Written Language

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Student Group | Kindergarten | 1 | 2 | 3–5 | 6–8 | 9–10 | 11 |
| All | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Male | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Female | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Asian | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Hispanic or Latino | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Special education services | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Economically disadvantaged | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Parent/Guardian education: some college or higher | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |

### Reference

Haberman, S. J. (1984). Adjustment by minimum discriminant information. *The Annals of Statistics,* *12*(3), 971–88.

## Score Comparability Results

This chapter provides results from analyses that evaluate the comparability of results observed for in-person and remote test takers.

### Item Performance Comparisons

The consistency of item performance was evaluated with three pieces of information: differential item functioning (DIF), classical item difficulty, and classical item discrimination.

### Differential Item Functioning

DIF is a commonly used approach for evaluating the consistency of item performance. It is largely used to identify differences suspected to be due to bias associated with gender or ethnicity. DIF does not require equivalent in-person and remote groups, and unweighted samples were used.

#### Dichotomous Items

The Mantel-Haenszel DIF (MH-DIF) statistic was calculated for dichotomous items (Mantel & Haenszel, 1959; Holland & Thayer, 1985). Using the total domain raw score as the criterion score, students in each domain score category in the focal group (e.g., remote test takers) are compared with examinees in the same score category in the reference group (e.g., in-person test takers).

Dichotomous items are assigned one of three DIF classifications shown in table 4.1. Differences are characterized as being large for items that are flagged with DIF at the C level, moderate for items flagged at the B level, and negligible for items flagged at the A level. For the purposes of this study, C level DIF was viewed as being an indicator of potential remote testing effects.

B and C level DIF are subdivided into B+ or B- and C+ and C- classifications. In this study, B+ or C+ classifications flagged items for which remote test takers performed better. B- and C- classifications flagged items for which in-person test takers performed better.

Table 4.1 DIF Categories for Dichotomous Items

|  |  |
| --- | --- |
| DIF category | Criteria |
| C (large) | Delta sub MH is at least 1.5 and is significantly greater than 1.0. |
| B (moderate) | Delta sub MHis at least 1.0 and is significantly greater than 0.0. |
| A (negligible) | Otherwise |

#### Polytomous Items

The standardization approach to DIF (Dorans & Schmitt, 1993; Zwick, Thayer, & Mazzeo, 1997; Dorans, 2013), in conjunction with the Mantel chi-square statistic (Mantel, 1963; Mantel & Haenszel, 1959), was used to identify polytomous items with DIF. The standard mean difference (SMD) compares the item means of the two groups after adjusting for differences in the distribution of students across the values of the matching variable, using the total domain raw score as the criterion score.

A positive SMDvalue meant that, conditional on the criterion score, the focal group (remote test takers) had a higher mean item score than the reference group (in-person test takers). In contrast, a negative SMD value means that, conditional upon the criterion score, the focal group had a lower mean item score than the reference group.

Polytomous items were assigned one of three DIF classifications shown in table 4.2. These classifications have the same interpretation as the dichotomous classifications.

Table 4.2 DIF Categories for Polytomous Items

|  |  |
| --- | --- |
| DIF category | Criteria |
| C (large) | * Mantel chi-square *p*-value is < .05; and * The absolute value of |SMD/SD| is > 0.25. |
| B (moderate) | * Mantel chi-square *p*-value is < 0.05; and * The absolute value is 0.17 < |SMD/SD| ≤ 0.25. |
| A (negligible) | * Mantel chi-square *p*-value is > .05; or * The absolute value of |SMD/SD| is ≤ 0.17. |

**Note:** SD = total group standard deviation of item score

These classifications were defined to be in alignment with the dichotomous classifications in terms of stringency (Zwick, Thayer, & Mazzeo, 1997).

#### Summary

Of the 419 items that were evaluated, only three were identified as having DIF at the C level. One item in kindergarten Listening had a C- DIF, indicating that in-person test takers performed better on that item than remote test takers. One item in grade two Speaking had a C+ DIF, indicating that remote test takers performed better on that item than in-person test takers. One item in grade two Reading had a C- DIF, indicating that in-person test takers performed better on that item than remote test takers. Counts of DIF classifications at each level are provided in table 4.3 for oral language and table 4.4 for written language. Bold values with asterisks are for C classifications.

Table 4.3 Differential Item Functioning Results: Oral Language

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Domain | DIF Category | Kindergarten | Grade 1 | Grade 2 | Grade Span 3–5 | Grade Span 6–8 | Grade Span 9–10 | Grade 11 |
| Listening | C- | **\*1** | 0 | 0 | 0 | 0 | 0 | 0 |
| Listening | B- | 3 | 2 | 0 | 0 | 0 | 0 | 0 |
| Listening | A | 16 | 20 | 22 | 22 | 21 | 22 | 22 |
| Listening | B+ | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| Listening | C+ | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Listening | Total | 20 | 22 | 22 | 22 | 22 | 22 | 22 |
| Speaking | C- | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Speaking | B- | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Speaking | A | 9 | 9 | 11 | 11 | 12 | 12 | 11 |
| Speaking | B+ | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| Speaking | C+ | 0 | 0 | **\*1** | 0 | 0 | 0 | 0 |
| Speaking | Total | 9 | 9 | 12 | 12 | 12 | 12 | 12 |

Table 4.4 Differential Item Functioning Results: Written Language

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Domain | DIF Category | Kindergarten | Grade 1 | Grade 2 | Grade Span 3–5 | Grade Span 6–8 | Grade Span 9–10 | Grade 11 |
| Reading | C- | 0 | 0 | **\*1** | 0 | 0 | 0 | 0 |
| Reading | B- | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| Reading | A | 13 | 20 | 23 | 26 | 26 | 26 | 24 |
| Reading | B+ | 1 | 1 | 1 | 0 | 0 | 0 | 2 |
| Reading | C+ | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Reading | Total | 14 | 21 | 26 | 26 | 26 | 26 | 26 |
| Writing | C- | N/A | N/A | N/A | 0 | 0 | 0 | 0 |
| Writing | B- | N/A | N/A | N/A | 0 | 0 | 0 | 0 |
| Writing | A | N/A | N/A | N/A | 6 | 6 | 6 | 6 |
| Writing | B+ | N/A | N/A | N/A | 0 | 0 | 0 | 0 |
| Writing | C+ | N/A | N/A | N/A | 0 | 0 | 0 | 0 |
| Writing | Total | N/A | N/A | N/A | 6 | 6 | 6 | 6 |

#### Item Difficulty

For dichotomous items, item difficulty is indicated by the *p*-value, which is the proportion of students who answer an item correctly. The range of *p*-values is from 0.00 to 1.00, inclusive. Items with higher *p*-values are easier items; those with lower *p*-values are more difficult items.

For polytomous items, difficulty is indicated by the average item score (AIS). The AIS can range from 0.00 to the maximum total possible points for an item. To facilitate interpretation, the AIS values for polytomous items or machine-scorable polytomous items are often expressed as the proportion of the maximum possible score, which is analogous to the *p‑*values of dichotomous items.

Table 4.5 contains the corresponding item difficulty information for the oral language items across grade levels and grade spans. It also contains the mean differences in those difficulty values as well as the smallest and largest differences observed.

The mean item difficulty differences for oral language are positive for some tests and negative for others with no obvious pattern. Minimum and maximum item difficulty differences are most extreme in the highest grade levels, where the item difficulties are calculated with the fewest test takers. Finally, numbers in table 4.5 that are bold and preceded with an asterisk (\*) indicate particularly low sample sizes.

Table 4.6 contains the corresponding item difficulty information for the written language items across grade levels and grade spans. Kindergarten mean differences correspond to more than a point difference in student raw scores. For higher grade levels or grade spans, mean differences were much smaller.

Numbers in table 4.6 that are bold and preceded with an asterisk (\*) indicate particularly low sample sizes.

Table 4.5 Item Difficulty Means: Oral Language

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | N Items | N In Person | Mean In Person | N Remote | Mean Remote | Min Difference | Max Difference | Mean Difference |
| Kindergarten | 29 | 819 | 0.82 | 797 | 0.84 | -0.02 | 0.09 | 0.03 |
| 1 | 31 | 1,201 | 0.86 | 2,035 | 0.85 | -0.07 | 0.03 | -0.01 |
| 2 | 34 | 1,522 | 0.89 | 2,265 | 0.88 | -0.09 | 0.01 | -0.02 |
| 3–5 | 34 | 3,329 | 0.77 | 5,492 | 0.77 | -0.04 | 0.04 | 0.01 |
| 6–8 | 34 | 1,583 | 0.78 | 3,077 | 0.76 | -0.12 | 0.03 | -0.02 |
| 9–10 | 34 | 514 | 0.73 | 1,317 | 0.72 | -0.13 | 0.08 | -0.01 |
| 11 | 34 | **\*154** | 0.75 | **\*380** | 0.75 | -0.16 | 0.16 | -0.01 |

**Note:** Values that are bolded and preceded with an asterisk are particularly small sample sizes.

Table 4.6 Item Difficulty Means: Written Language

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | N Items | N In Person | Mean In Person | N Remote | Mean Remote | Min Difference | Max Difference | Mean Difference |
| Kindergarten | 22 | **\*493** | 0.84 | **\*131** | 0.90 | -0.01 | 0.14 | 0.06 |
| 1 | 28 | 829 | 0.78 | **\*198** | 0.79 | -0.04 | 0.08 | 0.02 |
| 2 | 32 | 905 | 0.86 | **\*241** | 0.86 | -0.06 | 0.06 | 0.00 |
| 3–5 | 32 | 4,240 | 0.56 | 6,726 | 0.56 | -0.04 | 0.05 | 0.00 |
| 6–8 | 32 | 2,294 | 0.49 | 3,336 | 0.49 | -0.07 | 0.07 | -0.01 |
| 9–10 | 32 | 724 | 0.62 | 1,383 | 0.61 | -0.07 | 0.08 | -0.01 |
| 11 | 32 | **\*217** | 0.58 | **\*430** | 0.58 | -0.14 | 0.11 | 0.01 |

**Note:** Values that are bolded and preceded with an asterisk are particularly small sample sizes.

Figure 4.1 through figure 4.7 contain scatterplots of oral language *p*-values calculated for fall 2020 remote and in-person test takers. Data tables corresponding to these figures can be found in the [appendix](#_Appendix), in table A.1.

Each dot is an item. The solid line is the identity line indicating what perfect agreement would be. The dotted line is the best fit line, which reflects the best fitting line between the observed *p*-values for the two groups. The sample sizes for the two groups are included in the figure caption, with the first number indicating the number of remote test takers and the second the in-person test takers.

For kindergarten, in figure 4.1, the dashed best fit line is above the identity line, suggesting that remote test takers had higher *p*-values than in-person test takers for most of the items.

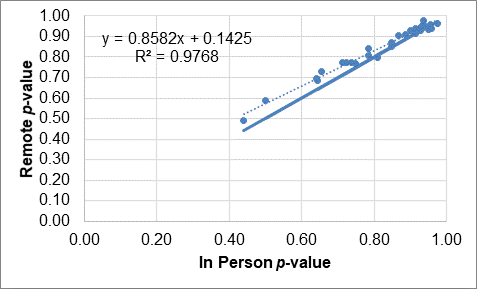


Figure 4.1 Item Difficulty Results: Oral Language, Kindergarten (N=797/819)

For grade one, in figure 4.2, the best fit line is nearly identical to the identity line, suggesting that remote test takers and in-person test takers were similar in performance.

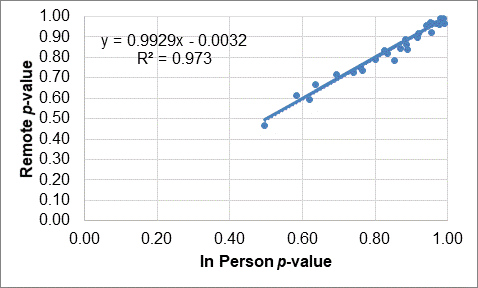


Figure 4.2 Item Difficulty Results: Oral Language, Grade One (N=2,035/1,201)

For grade two, in figure 4.3, the best fit line was below the identity line, suggesting that remote test takers had lower p-values than in-person test takers for most of the items.

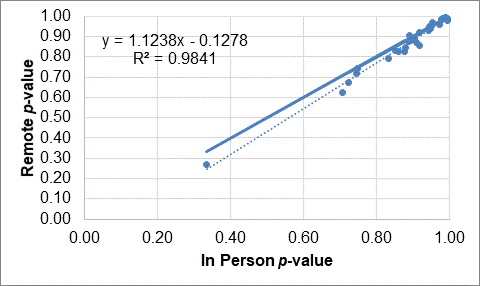


Figure 4.3 Item Difficulty Results: Oral Language, Grade Two (N=2,265/1,522)

For grade span three through five, in figure 4.4, the best fit line is nearly identical to the identity line, suggesting that remote test takers and in-person test takers were similar in performance.

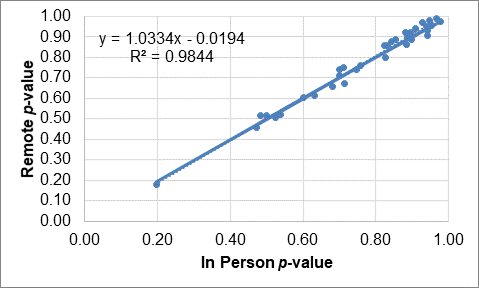


Figure 4.4 Item Difficulty Results: Oral Language, Grade Span Three Through Five (N=5,492/3,329)

For grade span six through eight, in figure 4.5, and grade span nine and ten, in figure 4.6, the best fit line is nearly identical to the identity line, suggesting that remote test takers and in-person test takers were similar in performance.

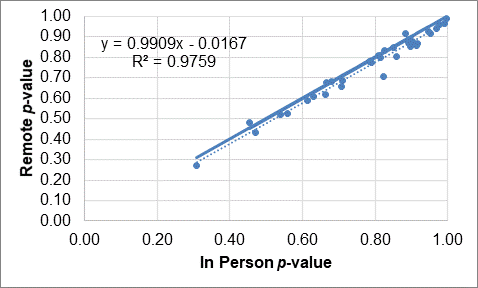


Figure 4.5 Item Difficulty Results: Oral Language, Grade Span Six Through Eight (N=3,077/1,583)

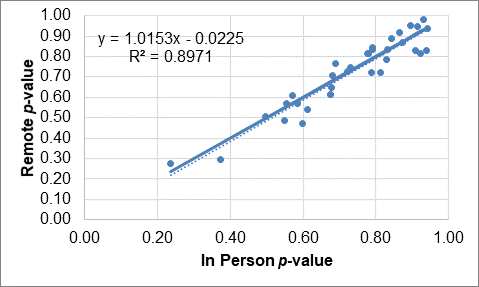


Figure 4.6 Item Difficulty Results: Oral Language, Grade Span Nine and Ten (N=1,317/514)

For grade eleven, in figure 4.7, the points for individual item *p*-values are fairly inconsistent, likely because fewer than 200 students participated in grade eleven.

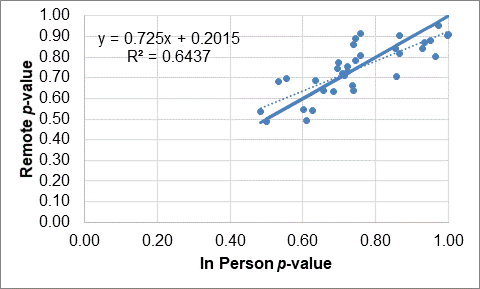


Figure 4.7 Item Difficulty Results: Oral Language, Grade Eleven (N=380/154)

Figure 4.8 through figure 4.14 contain scatterplots of written language *p*-values calculated for fall 2020 remote and in-person test takers. Data tables corresponding to these figures can be found in the [appendix](#_Appendix), in table A.2.

For kindergarten, in figure 4.8, the best fit line is above the identity line, suggesting that remote test takers had higher *p*-values than in-person test takers. Only 131 students responded to the remote version of the kindergarten test, which is a very small sample.

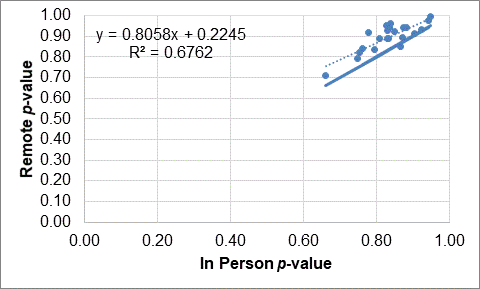


Figure 4.8 Item Difficulty Results: Written Language, Kindergarten (N=131/493)

For grade one, in figure 4.9, the best fit line is slightly above the identity line, suggesting that remote test takers had higher *p*-values than in-person test takers for most of the items.

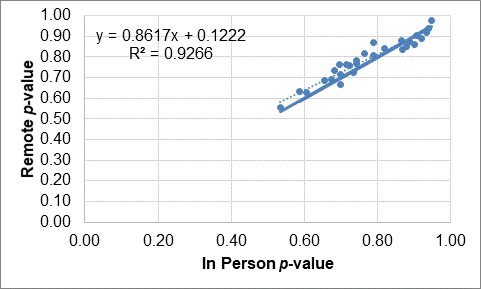


Figure 4.9 Item Difficulty Results: Written Language, Grade One (N=198/829)

The best fit line is nearly identical to the identity line, suggesting that remote test takers and in-person test takers were similar in performance for grade two, in figure 4.10; grade span three through five, in figure 4.11; grade span six through eight, in figure 4.12; grade span nine and ten, in figure 4.13; and grade eleven, in figure 4.14.

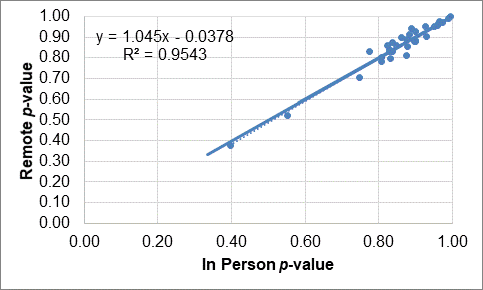


Figure 4.10 Item Difficulty Results: Written Language, Grade Two (N=241/905)

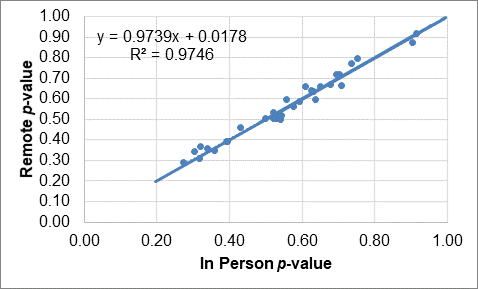


Figure 4.11 Item Difficulty Results: Written Language, Grade Span Three Through Five (N=6,726/4,240)

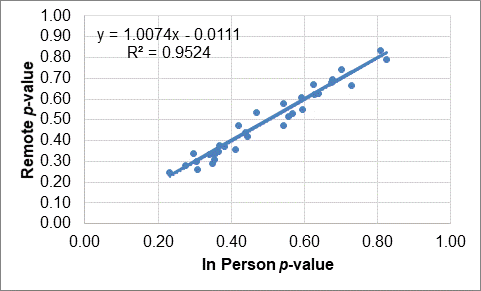


Figure 4.12 Item Difficulty Results: Written Language, Grade Span Six Through Eight (N=3,336/2,294)

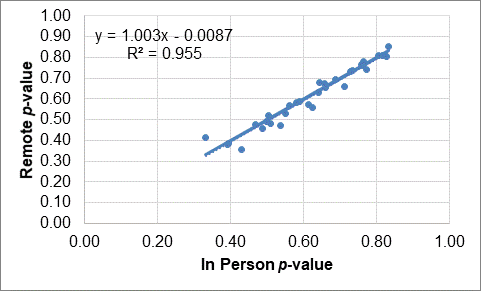


Figure 4.13 Item Difficulty Results: Written Language, Grade Span Nine and Ten (N=1,383/724)



Figure 4.14 Item Difficulty Results: Written Language, Grade Eleven (N=430/217)

#### Item-total Correlations

Item discrimination evaluates the degree to which individual items reflect the skill being measured by the test. Values typically range from zero to one, with larger values indicating that the item is a better measure of the targeted skill. Item discrimination was compared for students testing in person and remotely with point-biserial correlations for dichotomous items, which is the Pearson correlation between the item score and the operational domain score. The Pearson correlation between the item score and the operational domain score is also used for polytomous items. These are collectively described as item-total correlations.

Table 4.7 contains mean oral language in-person and remote item-total correlation values for in-person and remote test takers across grade levels and grade spans. It also contains the mean differences in those item-total correlation values as well as the smallest and largest differences observed. Mean differences ranged from -.02 to .06. Numbers in table 4.7 that are colored, bold, and preceded with an asterisk (\*) indicate particularly low sample sizes.

Table 4.8 contains the corresponding item-total correlation information for written language. Mean differences ranged from -.04 to .01. Numbers in table 4.8 that are bold and preceded with an asterisk (\*) indicate particularly low sample sizes.

Table 4.7 Item-Total Correlation Means: Oral Language

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | N Items | N In Person | Mean In Person | N Remote | Mean Remote | Min Difference | Max Difference | Mean Difference |
| Kindergarten | 29 | 819 | 0.58 | 797 | 0.56 | -0.22 | 0.07 | -0.02 |
| 1 | 31 | 1,201 | 0.37 | 2,035 | 0.39 | -0.12 | 0.14 | 0.02 |
| 2 | 34 | 1,522 | 0.36 | 2,265 | 0.34 | -0.24 | 0.16 | -0.02 |
| 3–5 | 34 | 3,329 | 0.44 | 5,492 | 0.41 | -0.27 | 0.12 | -0.03 |
| 6–8 | 34 | 1,583 | 0.42 | 3,077 | 0.49 | -0.19 | 0.42 | 0.06 |
| 9–10 | 34 | 514 | 0.53 | 1,317 | 0.50 | -0.44 | 0.22 | -0.03 |
| 11 | 34 | **\*154** | 0.47 | **\*380** | 0.53 | -0.43 | 0.74 | 0.06 |

Table 4.8 Item-Total Correlation Means: Written Language

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | N Items | N In Person | Mean In Person | N Remote | Mean Remote | Min Difference | Max Difference | Mean Difference |
| Kindergarten | 22 | **\*493** | 0.64 | **\*131** | 0.60 | -0.41 | 0.18 | -0.04 |
| 1 | 28 | 829 | 0.59 | **\*198** | 0.60 | -0.20 | 0.14 | 0.01 |
| 2 | 32 | 905 | 0.46 | **\*241** | 0.46 | -0.26 | 0.17 | -0.02 |
| 3–5 | 32 | 4,240 | 0.48 | 6,726 | 0.45 | -0.13 | 0.07 | -0.02 |
| 6–8 | 32 | 2,294 | 0.45 | 3,336 | 0.44 | -0.16 | 0.24 | -0.01 |
| 9–10 | 32 | 724 | 0.47 | 1,383 | 0.48 | -0.13 | 0.16 | 0.01 |
| 11 | 32 | **\*217** | 0.47 | **\*430** | 0.48 | -0.25 | 0.26 | 0.01 |

Figure 4.15 through figure 4.21 contain scatterplots of oral language item-total correlations calculated for fall 2020 remote and in-person test takers. Each dot in a scatterplot is an item. The solid line is the identity line indicating perfect agreement. The dotted line is the best fit line, which reflects the best fitting line between the observed item-total correlations. Because item-total correlations are sensitive to extreme *p*-values, these plots only contain items that had *p*-values less than .95 for both remote and in-person test takers. Data tables corresponding to these figures can be found in the [appendix](#_Appendix), in table A.1.

Figure 4.15 contains oral language item-total correlations for kindergarten students. Points are consistent with the exception of a single item; this item had a remote *p*-value of .94.



Figure 4.15 Item-Total Correlation Results: Oral Language, Kindergarten (N=797/819)

Figure 4.16 contains oral language item-total correlations for grade one students. Points are consistent with the exception of a few of the least discriminating items having slightly higher item-total correlations for remote test takers.

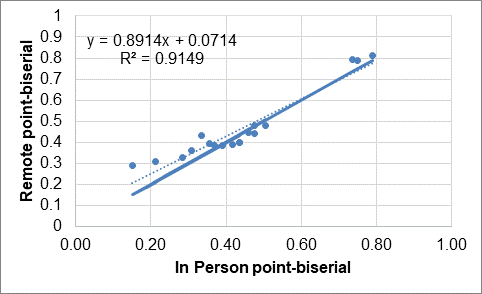


Figure 4.16 Item-Total Correlation Results: Oral Language, Grade One (N=2,035/1,201)

Figure 4.17 contains oral language item-total correlations for grade two students. Points are consistent, with identity and best fit lines being nearly identical.

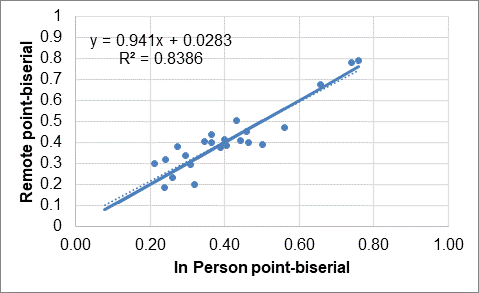


Figure 4.17 Item-Total Correlation Results: Oral Language, Grade Two (N=2,265/1,522)

Figure 4.18 contains oral language item-total correlations for grade span three through five students. Item-total correlations were fairly consistent, with the most discriminating items being a bit less discriminating for remote test takers.

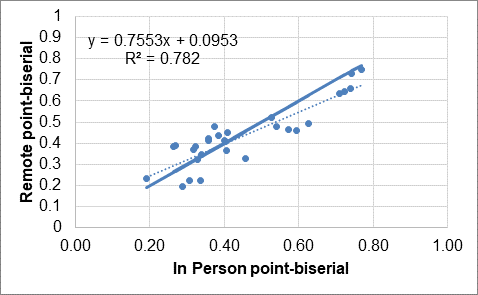


Figure 4.18 Item-Total Correlation Results: Oral Language, Grade Span Three Through Five (N=5,492/3,329)

Figure 4.19 contains oral language item-total correlations for grade span six through eight students. Item-total correlations were fairly consistent, with more items being slightly more discriminating for remote test takers than for in-person test takers.

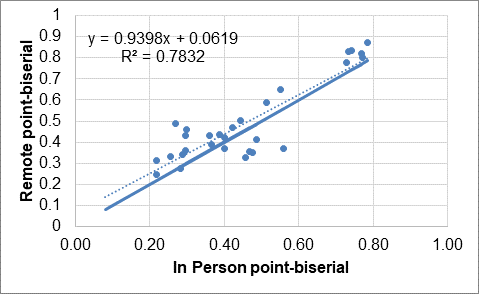


Figure 4.19 Item-Total Correlation Results: Oral Language, Grade Span Six Through Eight (N=3,077/1,583)

Figure 4.20 contains oral language item-total correlations for grade span nine and ten students. Item-total correlations were relatively less consistent for grade span nine and ten. The most discriminating items were slightly less discriminating for remote test takers.

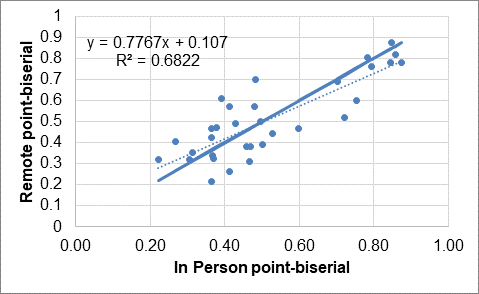


Figure 4.20 Item-Total Correlation Results: Oral Language, Grade Span Nine and Ten (N=1,317/514)

Figure 4.21 contains oral language item-total correlations for grade eleven students. Item-total correlations were least consistent for grade eleven, likely due to low sample sizes.

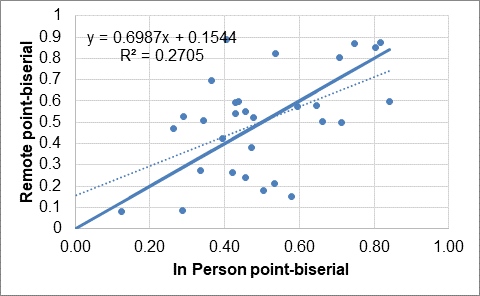


Figure 4.21 Item-Total Correlation Results: Oral Language, Grade Eleven (N=380/154)

Figure 4.22 through Figure 4.28 contain scatterplots of written language item-total correlations calculated for fall 2020 remote and in-person test takers. Data tables corresponding to these figures can be found in the [appendix](#_Appendix), in table A.2. The low sample sizes in kindergarten through grade two can be attributed to the nature of the writing administration, which requires paper–pencil materials to be returned for scoring.

Figure 4.22 contains written language item-total correlations for kindergarten students. Item-total correlations were fairly consistent, with more discriminating items being a bit less discriminating for remote test takers.

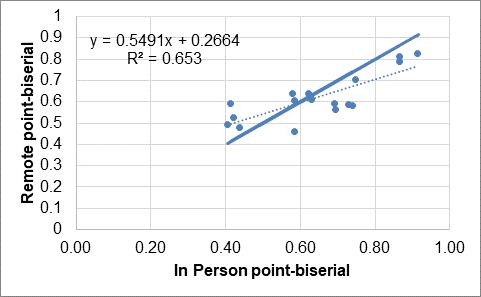


Figure 4.22 Item-Total Correlation Results: Written Language, Kindergarten (N=131/493)

Figure 4.23 contains written language item-total correlations for grade one students. Points are consistent, with identity and best fit lines being nearly identical.

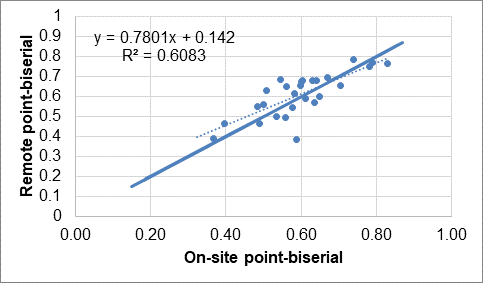


Figure 4.23 Item-Total Correlation Results: Written Language, Grade One (N=198/829)

Figure 4.24 contains written language item-total correlations for grade two students. Points are fairly consistent, with identity and best fit lines being nearly identical.

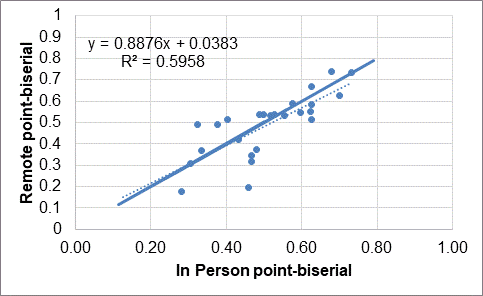


Figure 4.24 Item-Total Correlation Results: Written Language, Grade Two (N=241/905)

Figure 4.25 contains written language item-total correlations for grade span three through five students. Points are very consistent, with identity and best fit lines being nearly identical.

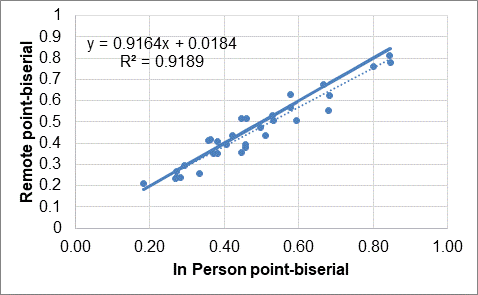


Figure 4.25 Item-Total Correlation Results: Written Language, Grade Span Three Through Five (N=6,726/4,240)

Figure 4.26 contains written language item-total correlations for grade span six through eight students. Item-total correlations were fairly consistent, with more discriminating items being a bit less discriminating for remote test takers. Best fit and identity lines are more closely aligned when the leftmost item in the graph, which has an item difficulty approaching chance-level performance, is removed.

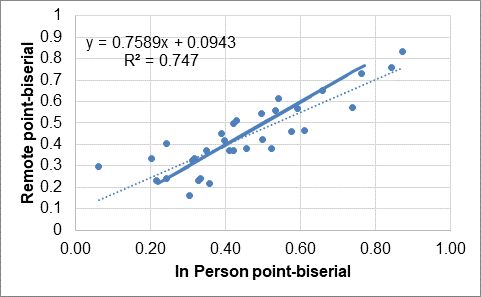


Figure 4.26 Item-Total Correlation Results: Written Language, Grade Span Six Through Eight (N=3,336/2,294)

Figure 4.27 contains written language item-total correlations for grade span nine and ten students. Item-total correlations were consistent, with more discriminating items being slightly less discriminating for remote test takers.

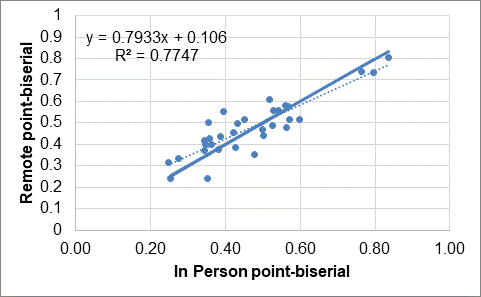


Figure 4.27 Item-Total Correlation Results: Written Language, Grade Span Nine and Ten (N=1,383/724)

Figure 4.28 contains written language item-total correlations for grade eleven students. Item-total correlations were fairly consistent, with points less consistent than observed in other grades or grade spans. Small sample sizes in grade eleven are a potential cause for this.

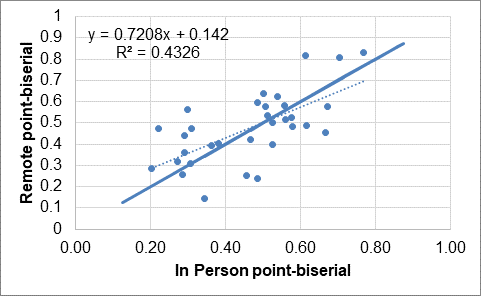


Figure 4.28 Item-Total Correlation Results: Written Language, Grade Eleven (N=430/217)

### Comparison of Score Means

Score means for in-person and remote students were compared using a weighted regression approach that was equivalent to performing a t-test for each grade or grade span. The comparison of mean scale scores for in-person and remote test takers is included for oral language in table 4.9. Based on a threshold *p*-value of .05, score means were significantly higher for remote kindergarten test takers and significantly lower for remote test takers in grade two. An asterisk appears next to those grade levels in the table. Based on Cohen’s thresholds (1988), an r value of .10 is judged to be a small effect size in the social sciences, which corresponds to an r-squared value of .01. Based on this threshold, the kindergarten mean difference is a small effect, and the grade two mean difference is very small.

Table 4.9 Scale Score Mean Differences: Oral Language Composite

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | In Person Mean | Remote Mean | Mean Difference (Remote–In Person) | F | Pr > F | r-squared |
| \*Kindergarten | 1471 | 1488 | 17.1 | 11.42 | <.01 | 0.014 |
| 1 | 1493 | 1492 | -0.6 | 0.08 | 0.77 | <0.001 |
| \*2 | 1530 | 1521 | -8.6 | 7.98 | <.01 | 0.007 |
| 3–5 | 1521 | 1525 | 4.0 | 1.86 | 0.17 | 0.001 |
| 6–8 | 1558 | 1550 | -8.2 | 1.51 | 0.22 | 0.002 |
| 9–10 | 1551 | 1560 | 8.9 | 0.51 | 0.48 | 0.001 |
| 11 | 1545 | 1551 | 5.7 | 0.04 | 0.84 | 0.001 |

The comparison of mean scale scores for in-person and remote test takers is included for written language in table 4.10. No significant differences were observed.

Table 4.10 Scale Score Mean Differences: Written Language Composite

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | In Person Mean | Remote Mean | Mean Difference (Remote–In Person) | F | Pr > F | r-squared |
| Kindergarten | 1491 | 1525 | 33.4 | 3.67 | 0.06 | 0.010 |
| 1 | 1480 | 1487 | 7.4 | 0.96 | 0.33 | 0.002 |
| 2 | 1522 | 1523 | 1.4 | 0.06 | 0.80 | <.001 |
| 3–5 | 1516 | 1516 | 0.3 | 0.01 | 0.93 | <.001 |
| 6–8 | 1536 | 1536 | 0.4 | <0.01 | 0.97 | <.001 |
| 9–10 | 1563 | 1565 | 1.6 | 0.17 | 0.68 | <.001 |
| 11 | 1558 | 1560 | 2.4 | 0.03 | 0.86 | <.001 |

### References

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge.

Dorans, N. J. (2013). *ETS contributions to the quantitative assessment of item, test, and score fairness*. (ETS Scientific and Policy Contributions Series SPC-13-04). Princeton, NJ: ETS.

Dorans, N., and Kulick, E. (1986). Demonstrating the utility of the standardization approach to assessing unexpected differential item performance on the Scholastic Aptitude Test. Journal of Educational Measurement, 23, 355–68.

Dorans, N. J., & Schmitt, A. P. (1993). Constructed response and differential item functioning: A pragmatic approach. In R. E. Bennett & W. C. Ward (Eds.), *Construction versus choice in cognitive measurement* (pp. 135–65). Erlbaum.

Holland, P. W., & Thayer, D. T. (1986). *Differential item functioning and the Mantel-Haenszel procedure* (ETS Research Report No. RR-86-31). Princeton, NJ: ETS.

Holland, P. W., & Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer & H. Braun (Eds.), *Test validity* (pp. 129–45). Erlbaum.

Mantel, N. (1963). Chi-Square tests with one degree of freedom: Extensions of the Mantel-Haenszel procedure. *Journal of the American Statistical Association*, *58*, 690–700.

Mantel, N. & Haenszel, W. (1959). Statistical aspects of the analyses of data from retrospective studies of disease. *Journal of the National Cancer Institute,* *22*, 719–48.

Zwick, R., Thayer, D. T., & Mazzeo, J. (1997). Descriptive and inferential procedures for assessing differential item functioning in polytomous items. *Applied Measurement in Education, 10*(4), 321–44.

## Supplemental Analysis Results

This chapter details results from analyses that provide additional information to support the evaluation of whether remote and in-person scores are comparable.

### Response Time

Response-time information was evaluated for students testing in person and remotely at each of the four ELPAC performance levels. Because the numbers of students in these classifications were very small in the lowest and highest grades, response-time information is provided only for students in grade spans three through five, six through eight, and nine and ten.

Oral language results for grade span three through five are provided in the box plots contained in figure 5.1. The figure contains box plots containing the mean, first quartile, second quartile, and third quartile for response time for remote and in-person students earning performance levels from one to four. The bottom of each box is the twenty-fifth percentile (also known as the first quartile or Q1), the top of each box is the seventy-fifth percentile (also known as the third quartile or Q3), the line through the box is the median, and the “+” symbol is the mean. Because student data is weighted, this information is based on weights rather than on the number of students. The figure shows that median response times were consistently longer for remote test takers. This difference is approximately 10 minutes for students in performance level two. This is an unfortunate outcome of the weighting of scores where a student with a long response time is assigned a large weight. For all the other performance levels, the differences in the medians were well under 10 minutes.

Figure 5.1 presents this data.

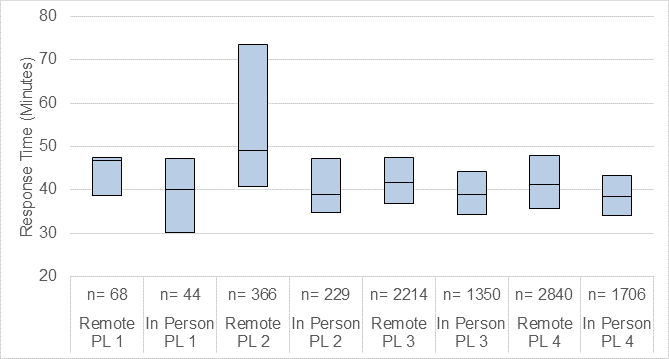


Figure 5.1 Response-Time Results: Oral Language, Grade Span Three Through Five

Data corresponding to figure 5.1 is provided in table 5.1.

Table 5.1 Response-Time Results: Oral Language, Grade Span Three Through Five

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Statistic | Remote PL1 | In Person PL1 | Remote PL2 | In Person PL2 | Remote PL3 | In Person PL3 | Remote PL4 | In Person PL4 |
| N | 68 | 44 | 366 | 229 | 2,214 | 1,350 | 2,840 | 1,706 |
| Mean | 43.84 | 38.7 | 58.08 | 42.03 | 43.46 | 40 | 42.61 | 39.91 |
| SD | 9.49 | 9.52 | 24.99 | 11.76 | 11.03 | 8.83 | 10.96 | 9.5 |
| Min | 15.14 | 20.87 | 9.45 | 13.96 | 16.33 | 13.99 | 4.19 | 9.86 |
| Max | 65.28 | 65.53 | 135.7 | 74.69 | 161.8 | 97.89 | 139.5 | 129.9 |
| Q1 | 38.78 | 30.06 | 40.75 | 34.79 | 36.74 | 34.41 | 35.59 | 34.15 |
| Q3 | 47.52 | 47.14 | 73.59 | 47.16 | 47.36 | 44.28 | 47.87 | 43.41 |
| Median | 46.77 | 39.99 | 48.98 | 38.95 | 41.59 | 38.96 | 41.16 | 38.37 |

Oral language response-time results for grade span six through eight are provided in the box plots contained in figure 5.2. Median response times differed by less than ten minutes for all performance levels. Median response times were shorter for remote test takers in performance level one and longer for remote test takers in performance levels two through four. Performance level one has only 70 students in in-person and remote groups combined, and that may contribute to the inconsistent differences.

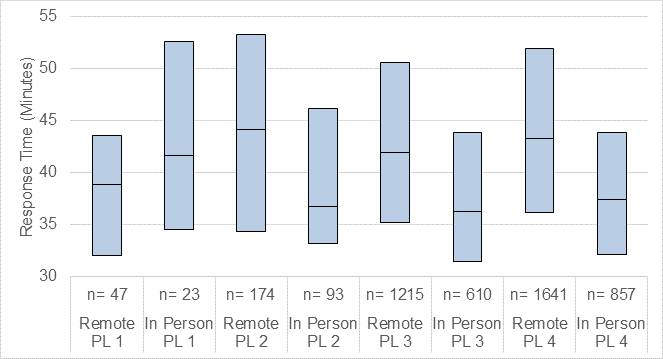


Figure 5.2 Response-Time Results: Oral Language, Grade Span Six Through Eight

Data corresponding to figure 5.2 is provided in table 5.2.

Table 5.2 Response-Time Results: Oral Language, Grade Span Six Through Eight

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Statistic | Remote PL1 | In Person PL1 | Remote PL2 | In Person PL2 | Remote PL3 | In Person PL3 | Remote PL4 | In Person PL4 |
| N | 47 | 23 | 174 | 93 | 1,215 | 610 | 1,641 | 857 |
| Mean | 38.10 | 43.08 | 43.98 | 40.22 | 44.49 | 39.19 | 46.19 | 40.43 |
| SD | 13.19 | 11.29 | 12.94 | 11.96 | 14.63 | 12.64 | 15.39 | 13.82 |
| Min | 11.32 | 18.53 | 17.63 | 1.66 | 5.68 | 19.57 | 12.86 | 12.54 |
| Max | 94.74 | 80.13 | 114.92 | 69.09 | 224.50 | 124.47 | 194.00 | 162.62 |
| Q1 | 32.05 | 34.57 | 34.30 | 33.20 | 35.17 | 31.46 | 36.13 | 32.17 |
| Q3 | 43.51 | 52.53 | 53.21 | 46.11 | 50.56 | 43.82 | 51.95 | 43.85 |
| Median | 38.82 | 41.62 | 44.09 | 36.74 | 41.96 | 36.23 | 43.23 | 37.42 |

Oral language response-time results for grade span nine and ten are provided in the box plots contained in figure 5.3. Median response-time differences were shorter for remote test takers in performance level one and longer for remote test takers in performance levels two through four. Performance level one has only 63 students in in-person and remote groups combined, so that may play a part in the inconsistent differences.

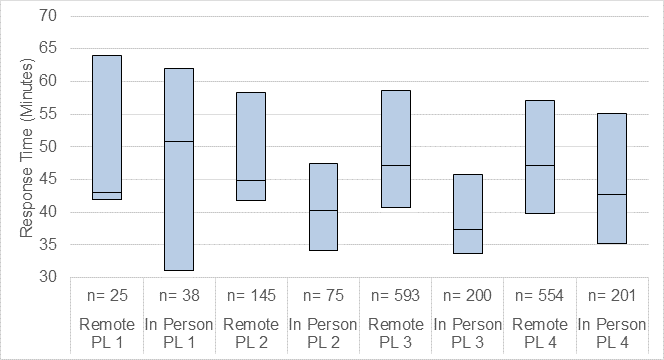


Figure 5.3 Response-Time Results: Oral Language, Grade Span Nine and Ten

Data corresponding to figure 5.3 is provided in table 5.3.

Table 5.3 Response-Time Results: Oral Language, Grade Span Nine and Ten

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Statistic | Remote PL1 | In Person PL1 | Remote PL2 | In Person PL2 | Remote PL3 | In Person PL3 | Remote PL4 | In Person PL4 |
| N | 25 | 38 | 145 | 75 | 593 | 200 | 554 | 201 |
| Mean | 52.41 | 47.49 | 49.51 | 42.77 | 50.95 | 41.27 | 51.18 | 47.80 |
| SD | 15.19 | 16.89 | 15.37 | 12.89 | 15.99 | 12.58 | 15.81 | 17.33 |
| Min | 19.25 | 15.74 | 15.85 | 20.30 | 20.03 | 15.56 | 16.68 | 19.92 |
| Max | 92.18 | 78.16 | 134.73 | 79.87 | 213.22 | 105.68 | 130.18 | 144.53 |
| Q1 | 41.99 | 31.11 | 41.72 | 34.22 | 40.71 | 33.63 | 39.79 | 35.15 |
| Q3 | 63.94 | 62.05 | 58.28 | 47.52 | 58.68 | 45.78 | 57.15 | 55.05 |
| Median | 43.07 | 50.85 | 44.86 | 40.22 | 47.14 | 37.41 | 47.12 | 42.65 |

Written language response-time results for grade span three through five are provided in the box plots contained in figure 5.4. Median response-time differences were longer for remote test takers in all performance levels.

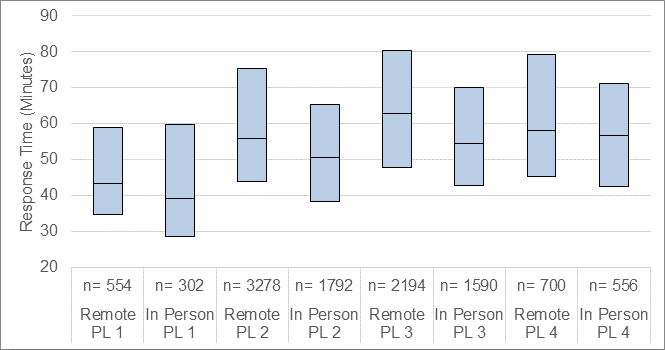


Figure 5.4 Response-Time Results: Written Language, Grade Span Three Through Five

Data corresponding to figure 5.4 is provided in table 5.4.

Table 5.4 Response Time Results: Written Language, Grade Span Three Through Five

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Statistic | Remote PL1 | In Person PL1 | Remote PL2 | In Person PL2 | Remote PL3 | In Person PL3 | Remote PL4 | In Person PL4 |
| N | 554 | 302 | 3,278 | 1,792 | 2,194 | 1,590 | 700 | 556 |
| Mean | 49.45 | 45.57 | 62.35 | 55.18 | 68.46 | 58.84 | 65.85 | 58.87 |
| SD | 23.76 | 24.97 | 29.00 | 24.89 | 29.83 | 23.35 | 29.65 | 22.81 |
| Min | 7.64 | 8.79 | 8.37 | 13.59 | 6.43 | 17.53 | 9.44 | 21.59 |
| Max | 182.33 | 166.57 | 515.64 | 323.91 | 251.65 | 212.66 | 236.10 | 174.03 |
| Q1 | 34.81 | 28.67 | 43.98 | 38.39 | 47.85 | 42.71 | 45.24 | 42.60 |
| Q3 | 58.82 | 59.81 | 75.40 | 65.47 | 80.44 | 69.98 | 79.19 | 71.16 |
| Median | 43.37 | 39.14 | 55.90 | 50.69 | 62.72 | 54.47 | 58.11 | 56.79 |

Written language response-time results for grade span six through eight are provided in the box plots contained in figure 5.5. Median response-time differences were longer for remote test takers in all performance levels. Response times for the first quartile and the median are similar for performance level one remote test takers, in part, because a student in this group was assigned a large weight.

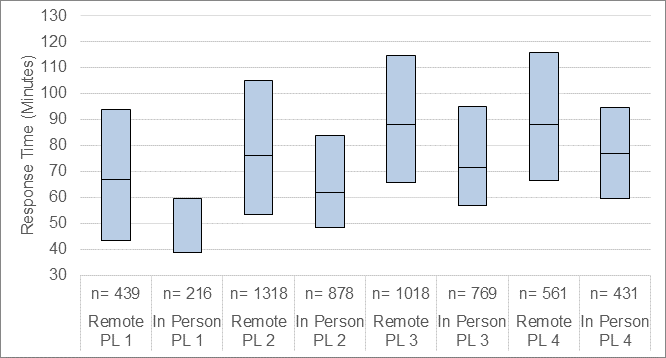


Figure 5.5 Response-Time Results: Written Language, Grade Span Six Through Eight

Data corresponding to figure 5.5 is provided in table 5.5.

Table 5.5 Response-Time Results: Written Language, Grade Span Six Through Eight

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Statistic | Remote PL1 | In Person PL1 | Remote PL2 | In Person PL2 | Remote PL3 | In Person PL3 | Remote PL4 | In Person PL4 |
| N | 439 | 216 | 1,318 | 878 | 1,018 | 769 | 561 | 431 |
| Mean | 71.55 | 52.09 | 86.88 | 71.94 | 95.90 | 78.51 | 99.81 | 85.32 |
| SD | 38.58 | 37.57 | 51.18 | 38.60 | 44.23 | 29.96 | 46.60 | 44.66 |
| Min | 1.58 | 12.87 | 21.25 | 6.86 | 22.30 | 23.30 | 18.40 | 24.12 |
| Max | 321.56 | 326.00 | 879.73 | 377.97 | 483.00 | 235.65 | 396.83 | 466.95 |
| Q1 | 43.32 | 38.81 | 53.55 | 48.27 | 65.80 | 56.96 | 66.48 | 59.60 |
| Q3 | 94.05 | 59.43 | 105.02 | 83.92 | 114.93 | 95.01 | 115.91 | 94.57 |
| Median | 66.80 | 38.84 | 76.19 | 61.73 | 88.13 | 71.71 | 87.99 | 77.07 |

Written language response-time results for grade span nine and ten are provided in the box plots contained in figure 5.6. Median response-time differences were longer for remote test takers in all performance levels.

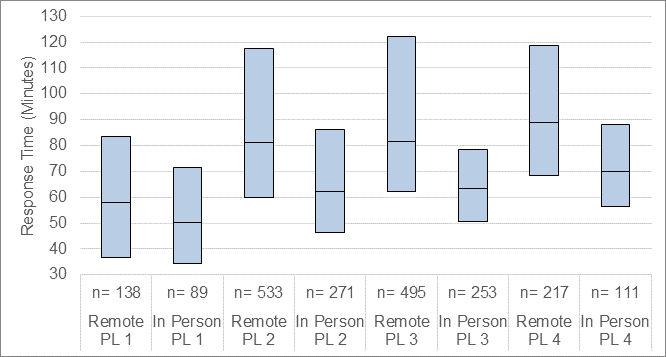


Figure 5.6 Response-Time Results: Written Language, Grade Span Nine and Ten

Data corresponding to figure 5.6 is provided in table 5.6.

Table 5.6 Response-Time Results: Written Language, Grade Span Nine and Ten

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Statistic | Remote PL1 | In Person PL1 | Remote PL2 | In Person PL2 | Remote PL3 | In Person PL3 | Remote PL4 | In Person PL4 |
| N | 138 | 89 | 533 | 271 | 495 | 253 | 217 | 111 |
| Mean | 65.09 | 56.19 | 98.96 | 68.25 | 97.49 | 66.31 | 101.08 | 79.00 |
| SD | 37.18 | 34.12 | 62.51 | 30.53 | 60.67 | 21.28 | 45.85 | 33.42 |
| Min | 4.85 | 11.75 | 19.45 | 20.89 | 27.40 | 23.20 | 33.27 | 29.83 |
| Max | 199.44 | 288.75 | 446.70 | 241.13 | 617.01 | 168.14 | 274.64 | 207.55 |
| Q1 | 36.57 | 34.30 | 59.82 | 46.39 | 62.38 | 50.44 | 68.38 | 56.47 |
| Q3 | 83.49 | 71.33 | 117.47 | 86.11 | 122.07 | 78.39 | 118.81 | 88.22 |
| Median | 57.83 | 50.39 | 81.00 | 62.11 | 81.73 | 63.53 | 88.70 | 69.79 |

### Score Reliability

The Summative ELPAC test reliabilities were evaluated for each domain using the coefficient alpha (Cronbach, 1951) index of internal consistency. Score reliabilities for each grade or grade span are provided for the Listening domain in table 5.7. Differences between remote and in person reliabilities were below .10 and varied in terms of whether remote or in person reliabilities were larger. Differences between remote and 2018–2019 reliabilities were uniformly negative, indicating that remote reliabilities were lower than 2018–2019 reliabilities. This is likely because there is less variance in the 2020 samples relative to the 2018–2019 population.

Table 5.7 Reliability: Listening

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | 2020 In Person | 2020 Remote | 2018–2019 | Difference (Remote–In Person) | Difference (Remote—2018–‍2019) |
| Kindergarten | 0.83 | 0.80 | 0.83 | -0.03 | -0.03 |
| 1 | 0.61 | 0.67 | 0.85 | 0.06 | -0.18 |
| 2 | 0.61 | 0.61 | 0.84 | <0.01 | -0.23 |
| 3–5 | 0.65 | 0.65 | 0.82 | <0.01 | -0.17 |
| 6–8 | 0.69 | 0.73 | 0.79 | 0.04 | -0.06 |
| 9–10 | 0.74 | 0.77 | 0.86 | 0.03 | -0.09 |
| 11 | 0.81 | 0.75 | 0.85 | -0.06 | -0.10 |

Score reliabilities for each grade or grade span are provided for the Speaking domain in table 5.8. Differences between remote and in person reliabilities were .13 for grade eleven. The remaining grades and grade spans were below .10 and varied in terms of whether remote or in person reliabilities were larger. Differences between remote and 2018–2019 reliabilities were uniformly negative; remote reliabilities were lower than 2018–2019 reliabilities.

Table 5.8 Reliability: Speaking

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | 2020 In Person | 2020 Remote | 2018–2019 | Difference (Remote–In Person) | Difference (Remote—2018–‍2019) |
| Kindergarten | 0.85 | 0.86 | 0.88 | 0.01 | -0.02 |
| 1 | 0.66 | 0.67 | 0.86 | 0.01 | -0.19 |
| 2 | 0.63 | 0.62 | 0.90 | -0.01 | -0.28 |
| 3–5 | 0.84 | 0.78 | 0.92 | -0.06 | -0.14 |
| 6–8 | 0.82 | 0.89 | 0.93 | 0.07 | -0.04 |
| 9–10 | 0.93 | 0.88 | 0.95 | -0.05 | -0.07 |
| 11 | 0.79 | 0.92 | 0.95 | 0.13 | -0.03 |

Score reliabilities for each grade or grade span are provided for the Reading domain in table 5.9. Differences between remote and in-person reliabilities had absolute values of .02 or below, with actual differences varying in terms of whether remote or in-person reliabilities were larger. Differences between remote and 2018–2019 reliabilities were predominantly negative; remote reliabilities tended to be lower than 2018–2019 reliabilities.

Table 5.9 Reliability: Reading

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | 2020 In Person | 2020 Remote | 2018–2019 | Difference (Remote–In Person) | Difference (Remote—2018–‍2019) |
| Kindergarten | 0.84 | 0.82 | 0.81 | -0.02 | 0.01 |
| 1 | 0.89 | 0.90 | 0.92 | 0.01 | -0.02 |
| 2 | 0.81 | 0.79 | 0.90 | -0.02 | -0.11 |
| 3–5 | 0.80 | 0.78 | 0.86 | -0.02 | -0.08 |
| 6–8 | 0.77 | 0.77 | 0.79 | <0.01 | -0.02 |
| 9–10 | 0.81 | 0.83 | 0.85 | 0.02 | -0.02 |
| 11 | 0.83 | 0.82 | 0.85 | -0.01 | -0.03 |

Score reliabilities for each grade or grade span are provided for the Writing domain in table 5.10. Differences between remote and in person reliabilities were .16 for grade eleven. The remaining grades and grade spans had differences that were negative. Differences between remote and 2018–2019 reliabilities were uniformly negative; remote reliabilities were lower than 2018–2019 reliabilities.

Table 5.10 Reliability: Writing

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | 2020 In Person | 2020 Remote | 2018–2019 | Difference (Remote–In Person) | Difference (Remote—2018–‍2019) |
| Kindergarten | 0.90 | 0.84 | 0.91 | -0.06 | -0.07 |
| 1 | 0.81 | 0.79 | 0.84 | -0.02 | -0.05 |
| 2 | 0.73 | 0.71 | 0.85 | -0.02 | -0.14 |
| 3–5 | 0.84 | 0.79 | 0.85 | -0.05 | -0.06 |
| 6–8 | 0.84 | 0.76 | 0.84 | -0.08 | -0.08 |
| 9–10 | 0.77 | 0.74 | 0.86 | -0.03 | -0.12 |
| 11 | 0.63 | 0.79 | 0.88 | 0.16 | -0.09 |

### Correlations Among Domain Scores

Correlations among domain scores for oral language are provided in table 5.11. Differences between remote and in person correlations were varied across grade levels and grade spans, with both positive and negative values observed with no apparent pattern. Differences between remote and 2018–2019 correlations were nearly all negative.

Table 5.11 Correlations Among Domain Raw Scores: Oral Language

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | 2020 In Person | 2020 Remote | 2018–2019 | Difference (Remote–In Person) | Difference (Remote—2018–‍2019) |
| Kindergarten | 0.67 | 0.67 | 0.67 | <0.01 | <0.01 |
| 1 | 0.40 | 0.32 | 0.69 | -0.08 | -0.37 |
| 2 | 0.23 | 0.23 | 0.67 | <0.01 | -0.44 |
| 3–5 | 0.44 | 0.30 | 0.64 | -0.14 | -0.34 |
| 6–8 | 0.49 | 0.58 | 0.65 | 0.11 | -0.07 |
| 9–10 | 0.56 | 0.63 | 0.70 | 0.07 | -0.07 |
| 11 | 0.44 | 0.44 | 0.68 | <0.01 | -0.24 |

Relatively low correlations for grade two oral language were associated with a restriction of range for the in-person and remote students as compared to the 2018–2019 students. A bubble plot of the 2018–2019 raw scores earned in the Listening and Speaking domains is provided in figure 5.7. Each point is a combination of raw scores on the Listening and Speaking domains. The size of the point reflects the number of students earning that combination of raw scores. Figure 5.7 shows positive linear relationship with raw scores of zero in both domains. Data for this bubble plot is provided in table 5.11.

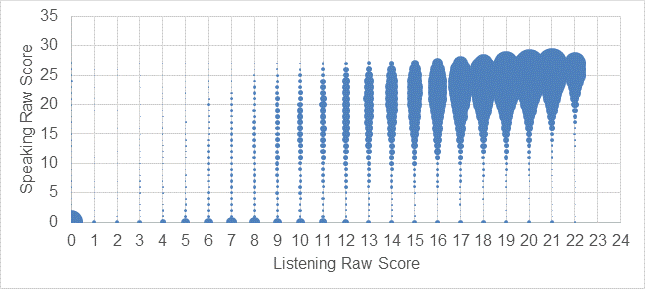


Figure 5.7 Follow-up for Correlations: Oral Language, Grade Two 2018–2019

Corresponding bubble plots of the raw scores earned in the Listening and Speaking domains for in-person test takers is provided in figure 5.8, and for remote test takers in figure 5.9.

Figure 5.8, for in-person test takers, shows positive linear relationship with scores downs to raw scores of zero in both domains, with a modest, positive, linear relationship with nearly all scores in top half of two raw score scales. Data for this bubble plot is provided in table 5.11.

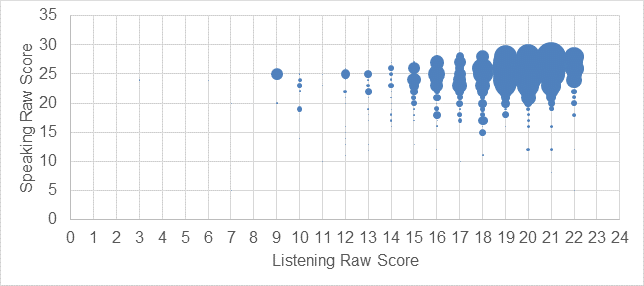


Figure 5.8 Follow-up for Correlations: Oral Language, Grade Two In Person

Figure 5.9, for remote test takers, shows a modest, positive, linear relationship with nearly all scores in top half of two raw score scales. Data for this bubble plot is provided in table 5.11.

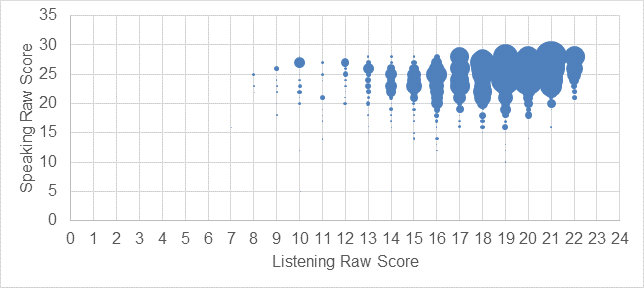


Figure 5.9 Follow-up for Correlations: Oral Language, Grade Two Remote

Correlations among domain scores for written language are provided in table 5.12. Differences between remote and in-person correlations varied across grade levels and grade spans, with both positive and negative values observed with no apparent pattern. Differences between remote and 2018–2019 correlations were nearly all negative.

Table 5.12 Correlations Among Domain Raw Scores: Written Language

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | 2020 In Person | 2020 Remote | 2018–2019 | Difference (Remote–In Person) | Difference (Remote—2018–‍2019) |
| Kindergarten | 0.70 | **\*0.13** | 0.69 | -0.57 | -0.56 |
| 1 | 0.65 | 0.72 | 0.76 | 0.07 | -0.04 |
| 2 | 0.43 | 0.50 | 0.76 | 0.07 | -0.26 |
| 3–5 | 0.71 | 0.63 | 0.74 | -0.08 | -0.11 |
| 6–8 | 0.66 | 0.55 | 0.60 | -0.11 | -0.05 |
| 9–10 | 0.57 | 0.61 | 0.67 | 0.04 | -0.06 |
| 11 | 0.62 | 0.54 | 0.70 | -0.08 | -0.16 |

The kindergarten written language difference between remote and in person raw score correlations was -0.57, which is a very large difference. The corresponding remote correlation, which is relatively small, is set in bold and is preceded by an asterisk, to highlight this unusual value.

Figure 5.10 is a bubble plot of the Reading domain and Writing domain raw scores for the in-person test takers. Each bubble is a combination of raw scores for the Reading and Writing domains. The size of the bubbles reflects the number of test takers earning that combination of raw scores. Many in-person students earned the maximum Writing score, and many of those students earned the maximum possible Reading score. The data table corresponding to this figure can be found in the [appendix](#_Appendix), in table A.3.

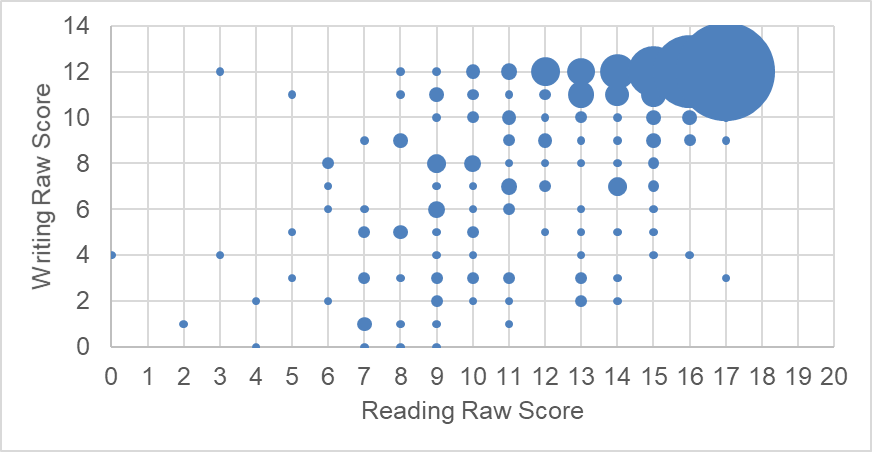


Figure 5.10 Follow-up for Correlations: Written Language, Kindergarten In Person

Figure 5.11 is a corresponding bubble plot of the Reading domain and Writing domain raw scores for the remote test takers. Many remote students earned the maximum Writing score, and many of those students earned the maximum possible Reading score. However, there were a few students earning values in the bottom half of either raw score range. The data table corresponding to this figure can be found in the [appendix](#_Appendix), in table A.4.

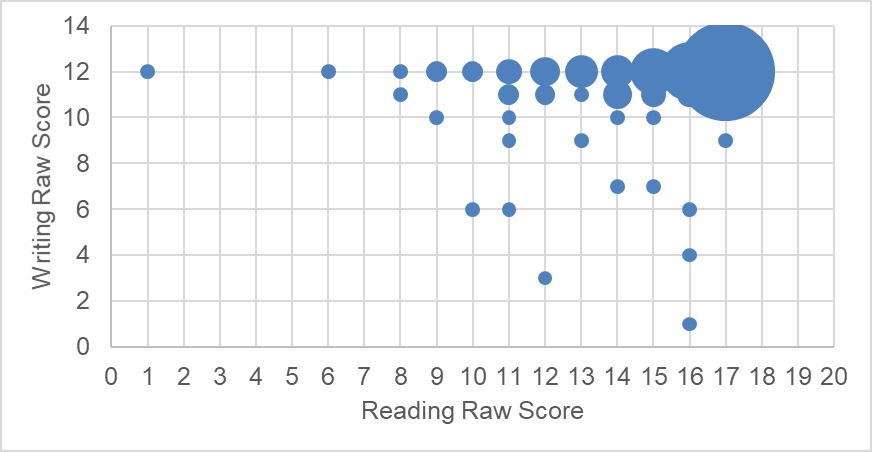


Figure 5.11 Follow-up for Correlations: Written Language, Kindergarten Remote

Having the full range of scores is important in getting an accurate measure of correlation. When the range is restricted for some reason, the correlation can be underestimated. That appears to be the case here for remote test takers. Reports came in from the field suggesting that schools were uncomfortable sending secure test materials home. Guidance was provided to the field to test students who were likely to be reclassified. It is possible that materials were only sent home to remote students who were expected to be reclassified.

## Conclusion

The results of the evaluation support the conclusion that remote testing can be viewed as reasonably comparable to in-person testing for the Summative ELPAC under the unusual conditions and with the atypical types of student samples observed here.

Comparisons of in person and remote ELPAC test administration results found a lot of consistency across the three item-based measures that were evaluated. DIF found only three item differences appearing in a seemingly random pattern across grade levels or grade spans and domains. Item difficulties were also very similar in nearly all tests. Kindergarten results were different by location; however, those differences may be the result of limitations of the matching approach for this small population of students. In grade two, small differences in item difficulty were also found. Item discrimination results were fairly consistent, with differences appearing when the numbers of test takers participating in a particular grade level or grade span were small.

Comparisons of test scores found differences in two of the fourteen ELPAC tests. Kindergarten oral language scores were slightly higher for remote test takers, which is consistent with the item difficulty differences and the theory that limitations in matching may have been at work. Grade two oral language scores were slightly lower for remote test takers; that difference accounted for less than 1 percent of the variance in student scores.

The validity and reliability analyses suggested no threats to the psychometric properties of the remotely administered tests. Response times were a bit longer for remote test takers, but nothing in these differences suggested anything problematic. Reliability estimates were similar for in-person and remote test taker scores; differences between them were associated with differences in score variability, suggesting that reduced variability was responsible for the differences. Differences in correlations between domains were present but followed no pattern that might have suggested a problem in what ELPAC tests measure across the two modes of test administration.

The results of this study support the conclusion that scores from remote testing can be viewed as reasonably comparable to scores from in-person testing for the Summative ELPAC under the unusual conditions and with the unusual types of student samples observed here. Further evidence should be gathered before more general and definitive judgments about comparability for the Summative ELPAC or other test populations are made.

## Appendix

This appendix contains data tables that include item statistics corresponding to figure 4.1 to figure 4.28. Table A.1 includes oral language *p*-values for in-person and remote test takers as well as the differences between them; it also contains point-biserial correlations for in-person and remote test takers as well as differences between them. Table A.2 provides the same information for written language.

Also included are data tables that contain raw scores corresponding to figure 5.10 and Figure 5.11. Table A.3 contains the data for figure 5.10; table A.4 contains the data for Figure 5.11.

Table A.1 Oral Language Classical Item Statistics

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level | Domain | Task Type | In Person *P*-value | Remote *P*-value | Remote–In Person *P*-value | In Person PtBis | Remote PtBis | Remote–In Person PtBis |
| Kindergarten | Listening | Listen to a Short Exchange | 0.75 | 0.77 | 0.02 | 0.49 | 0.55 | 0.06 |
| Kindergarten | Listening | Listen to a Short Exchange | 0.95 | 0.93 | -0.01 | 0.48 | 0.54 | 0.05 |
| Kindergarten | Listening | Listen to a Short Exchange | 0.93 | 0.93 | 0.00 | 0.53 | 0.56 | 0.03 |
| Kindergarten | Listening | Listen to a Short Exchange | 0.93 | 0.94 | 0.02 | 0.55 | 0.45 | -0.10 |
| Kindergarten | Listening | Listen to a Short Exchange | 0.97 | 0.97 | -0.01 | 0.53 | 0.46 | -0.07 |
| Kindergarten | Listening | Listen to a Story | 0.91 | 0.94 | 0.02 | 0.54 | 0.33 | -0.22 |
| Kindergarten | Listening | Listen to a Story | 0.85 | 0.85 | 0.01 | 0.50 | 0.49 | -0.01 |
| Kindergarten | Listening | Listen to a Story | 0.78 | 0.84 | 0.06 | 0.65 | 0.60 | -0.05 |
| Kindergarten | Listening | Listen to a Story | 0.94 | 0.98 | 0.04 | 0.60 | 0.44 | -0.16 |
| Kindergarten | Listening | Listen to a Story | 0.50 | 0.59 | 0.09 | 0.41 | 0.46 | 0.05 |
| Kindergarten | Listening | Listen to a Story | 0.90 | 0.93 | 0.03 | 0.50 | 0.57 | 0.07 |
| Kindergarten | Listening | Listen to a Story | 0.93 | 0.95 | 0.02 | 0.59 | 0.40 | -0.19 |
| Kindergarten | Listening | Listen to a Story | 0.81 | 0.80 | -0.01 | 0.57 | 0.45 | -0.11 |
| Kindergarten | Listening | Listen to a Story | 0.79 | 0.81 | 0.02 | 0.52 | 0.51 | -0.01 |
| Kindergarten | Listening | Listen to an Oral Presentation | 0.44 | 0.49 | 0.05 | 0.43 | 0.50 | 0.07 |
| Kindergarten | Listening | Listen to an Oral Presentation | 0.66 | 0.73 | 0.08 | 0.49 | 0.50 | 0.01 |
| Kindergarten | Listening | Listen to an Oral Presentation | 0.96 | 0.94 | -0.02 | 0.56 | 0.48 | -0.08 |
| Kindergarten | Listening | Listen to an Oral Presentation | 0.71 | 0.78 | 0.06 | 0.40 | 0.39 | -0.01 |
| Kindergarten | Listening | Listen to an Oral Presentation | 0.74 | 0.77 | 0.04 | 0.44 | 0.45 | 0.01 |
| Kindergarten | Listening | Listen to an Oral Presentation | 0.85 | 0.86 | 0.01 | 0.53 | 0.45 | -0.08 |
| Kindergarten | Speaking | Talk About a Scene | 0.96 | 0.96 | 0.00 | 0.53 | 0.59 | 0.06 |
| Kindergarten | Speaking | Talk About a Scene | 0.94 | 0.96 | 0.03 | 0.54 | 0.45 | -0.10 |
| Kindergarten | Speaking | Talk About a Scene | 0.91 | 0.92 | 0.00 | 0.71 | 0.70 | -0.01 |
| Kindergarten | Speaking | Talk About a Scene | 0.89 | 0.91 | 0.02 | 0.71 | 0.75 | 0.04 |
| Kindergarten | Speaking | Support an Opinion | 0.87 | 0.91 | 0.04 | 0.67 | 0.75 | 0.07 |

Table A.1 *(continuation one)*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level | Domain | Task Type | In Person *P*-value | Remote *P*-value | Remote–In Person *P*-value | In Person PtBis | Remote PtBis | Remote–In Person PtBis |
| Kindergarten | Speaking | Support an Opinion | 0.85 | 0.87 | 0.02 | 0.70 | 0.77 | 0.07 |
| Kindergarten | Speaking | Retell a Narrative | 0.64 | 0.68 | 0.04 | 0.85 | 0.88 | 0.03 |
| Kindergarten | Speaking | Retell a Narrative | 0.72 | 0.77 | 0.05 | 0.87 | 0.89 | 0.02 |
| Kindergarten | Speaking | Summarize an Academic Presentation | 0.64 | 0.69 | 0.05 | 0.81 | 0.83 | 0.02 |
| 1 | Listening | Listen to a Short Exchange | 0.76 | 0.73 | -0.03 | 0.21 | 0.31 | 0.10 |
| 1 | Listening | Listen to a Short Exchange | 0.85 | 0.79 | -0.07 | 0.15 | 0.29 | 0.14 |
| 1 | Listening | Listen to a Classroom Conversation | 0.62 | 0.59 | -0.03 | 0.36 | 0.39 | 0.04 |
| 1 | Listening | Listen to a Classroom Conversation | 0.89 | 0.84 | -0.05 | 0.29 | 0.33 | 0.04 |
| 1 | Listening | Listen to a Story | 0.99 | 0.97 | -0.02 | 0.18 | 0.23 | 0.05 |
| 1 | Listening | Listen to a Story | 0.99 | 0.99 | 0.00 | 0.21 | 0.18 | -0.03 |
| 1 | Listening | Listen to a Story | 0.83 | 0.82 | -0.01 | 0.43 | 0.40 | -0.03 |
| 1 | Listening | Listen to a Story | 0.99 | 0.97 | -0.02 | 0.21 | 0.27 | 0.06 |
| 1 | Listening | Listen to a Story | 0.98 | 0.96 | -0.02 | 0.28 | 0.35 | 0.06 |
| 1 | Listening | Listen to a Story | 0.89 | 0.86 | -0.02 | 0.37 | 0.38 | 0.01 |
| 1 | Listening | Listen to a Story | 0.98 | 0.97 | -0.01 | 0.26 | 0.32 | 0.06 |
| 1 | Listening | Listen to a Story | 0.98 | 0.97 | -0.01 | 0.19 | 0.20 | 0.01 |
| 1 | Listening | Listen to a Story | 0.95 | 0.92 | -0.03 | 0.24 | 0.37 | 0.13 |
| 1 | Listening | Listen to an Oral Presentation | 0.92 | 0.90 | -0.02 | 0.42 | 0.39 | -0.03 |
| 1 | Listening | Listen to an Oral Presentation | 0.64 | 0.67 | 0.03 | 0.51 | 0.48 | -0.02 |
| 1 | Listening | Listen to an Oral Presentation | 0.87 | 0.85 | -0.03 | 0.46 | 0.45 | -0.01 |
| 1 | Listening | Listen to an Oral Presentation | 0.88 | 0.89 | 0.00 | 0.34 | 0.43 | 0.10 |
| 1 | Listening | Listen to an Oral Presentation | 0.80 | 0.79 | -0.02 | 0.31 | 0.36 | 0.05 |
| 1 | Listening | Listen to an Oral Presentation | 0.92 | 0.92 | 0.00 | 0.39 | 0.39 | 0.00 |
| 1 | Listening | Listen to an Oral Presentation | 0.50 | 0.47 | -0.03 | 0.48 | 0.48 | 0.00 |

Table A.1 *(continuation two)*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level | Domain | Task Type | In Person *P*-value | Remote *P*-value | Remote–In Person *P*-value | In Person PtBis | Remote PtBis | Remote–In Person PtBis |
| 1 | Listening | Listen to an Oral Presentation | 0.69 | 0.72 | 0.03 | 0.43 | 0.40 | -0.03 |
| 1 | Listening | Listen to an Oral Presentation | 0.74 | 0.72 | -0.02 | 0.48 | 0.44 | -0.03 |
| 1 | Speaking | Talk About a Scene | 0.98 | 0.99 | 0.01 | 0.26 | 0.19 | -0.07 |
| 1 | Speaking | Talk About a Scene | 0.95 | 0.97 | 0.02 | 0.30 | 0.18 | -0.12 |
| 1 | Speaking | Talk About a Scene | 0.94 | 0.95 | 0.01 | 0.32 | 0.39 | 0.07 |
| 1 | Speaking | Talk About a Scene | 0.96 | 0.96 | 0.01 | 0.39 | 0.44 | 0.05 |
| 1 | Speaking | Support an Opinion | 0.95 | 0.96 | 0.01 | 0.46 | 0.37 | -0.09 |
| 1 | Speaking | Support an Opinion | 0.97 | 0.97 | 0.00 | 0.36 | 0.31 | -0.05 |
| 1 | Speaking | Retell a Narrative | 0.83 | 0.83 | 0.01 | 0.75 | 0.79 | 0.04 |
| 1 | Speaking | Summarize an Academic Presentation | 0.76 | 0.75 | -0.01 | 0.79 | 0.81 | 0.02 |
| 1 | Speaking | Summarize an Academic Presentation | 0.58 | 0.61 | 0.03 | 0.74 | 0.79 | 0.06 |
| 2 | Listening | Listen to a Short Exchange | 0.99 | 0.99 | 0.00 | 0.11 | 0.08 | -0.03 |
| 2 | Listening | Listen to a Short Exchange | 0.90 | 0.89 | -0.02 | 0.24 | 0.32 | 0.08 |
| 2 | Listening | Listen to a Classroom Conversation | 0.91 | 0.87 | -0.04 | 0.27 | 0.38 | 0.11 |
| 2 | Listening | Listen to a Classroom Conversation | 0.88 | 0.85 | -0.04 | 0.35 | 0.40 | 0.06 |
| 2 | Listening | Listen to a Story | 1.00 | 0.98 | -0.02 | 0.11 | 0.27 | 0.16 |
| 2 | Listening | Listen to a Story | 0.89 | 0.88 | -0.02 | 0.50 | 0.39 | -0.11 |
| 2 | Listening | Listen to a Story | 0.94 | 0.93 | -0.01 | 0.46 | 0.45 | -0.01 |
| 2 | Listening | Listen to a Story | 0.92 | 0.86 | -0.06 | 0.46 | 0.40 | -0.07 |
| 2 | Listening | Listen to a Story | 0.97 | 0.96 | -0.02 | 0.33 | 0.19 | -0.14 |
| 2 | Listening | Listen to a Story | 0.95 | 0.94 | -0.01 | 0.31 | 0.29 | -0.01 |
| 2 | Listening | Listen to a Story | 0.98 | 0.98 | 0.00 | 0.36 | 0.26 | -0.10 |
| 2 | Listening | Listen to a Story | 0.98 | 0.98 | 0.00 | 0.42 | 0.18 | -0.24 |

Table A.1 *(continuation three)*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level | Domain | Task Type | In Person *P*-value | Remote *P*-value | Remote–In Person *P*-value | In Person PtBis | Remote PtBis | Remote–In Person PtBis |
| 2 | Listening | Listen to a Story | 0.92 | 0.92 | 0.00 | 0.26 | 0.23 | -0.03 |
| 2 | Listening | Listen to an Oral Presentation | 0.83 | 0.79 | -0.04 | 0.36 | 0.40 | 0.03 |
| 2 | Listening | Listen to an Oral Presentation | 0.86 | 0.83 | -0.04 | 0.30 | 0.34 | 0.04 |
| 2 | Listening | Listen to an Oral Presentation | 0.71 | 0.62 | -0.09 | 0.37 | 0.44 | 0.07 |
| 2 | Listening | Listen to an Oral Presentation | 0.33 | 0.27 | -0.06 | 0.32 | 0.20 | -0.12 |
| 2 | Listening | Listen to an Oral Presentation | 0.88 | 0.83 | -0.05 | 0.43 | 0.51 | 0.07 |
| 2 | Listening | Listen to an Oral Presentation | 0.85 | 0.83 | -0.02 | 0.56 | 0.47 | -0.09 |
| 2 | Listening | Listen to an Oral Presentation | 0.89 | 0.87 | -0.01 | 0.21 | 0.30 | 0.09 |
| 2 | Listening | Listen to an Oral Presentation | 0.72 | 0.67 | -0.05 | 0.40 | 0.41 | 0.01 |
| 2 | Listening | Listen to an Oral Presentation | 0.95 | 0.95 | 0.00 | 0.24 | 0.18 | -0.05 |
| 2 | Speaking | Talk About a Scene | 0.99 | 0.99 | 0.00 | 0.19 | 0.05 | -0.14 |
| 2 | Speaking | Talk About a Scene | 0.99 | 1.00 | 0.01 | 0.08 | 0.05 | -0.02 |
| 2 | Speaking | Talk About a Scene | 0.98 | 0.99 | 0.01 | 0.26 | 0.14 | -0.12 |
| 2 | Speaking | Talk About a Scene | 0.99 | 0.99 | 0.00 | 0.17 | 0.17 | -0.01 |
| 2 | Speaking | Speech Functions | 0.95 | 0.94 | -0.01 | 0.40 | 0.38 | -0.02 |
| 2 | Speaking | Speech Functions | 0.89 | 0.91 | 0.01 | 0.44 | 0.41 | -0.03 |
| 2 | Speaking | Speech Functions | 0.95 | 0.95 | 0.00 | 0.39 | 0.37 | -0.02 |
| 2 | Speaking | Support an Opinion | 0.99 | 0.99 | 0.00 | 0.26 | 0.19 | -0.08 |
| 2 | Speaking | Support an Opinion | 0.96 | 0.97 | 0.01 | 0.42 | 0.32 | -0.10 |
| 2 | Speaking | Retell a Narrative | 0.90 | 0.88 | -0.01 | 0.66 | 0.68 | 0.02 |
| 2 | Speaking | Summarize an Academic Presentation | 0.75 | 0.72 | -0.03 | 0.76 | 0.79 | 0.03 |
| 2 | Speaking | Summarize an Academic Presentation | 0.75 | 0.74 | -0.01 | 0.74 | 0.78 | 0.04 |
| 3 | Listening | Listen to a Short Exchange | 0.83 | 0.80 | -0.03 | 0.26 | 0.38 | 0.12 |
| 3 | Listening | Listen to a Short Exchange | 0.89 | 0.90 | 0.01 | 0.34 | 0.22 | -0.11 |

Table A.1 *(continuation four)*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level | Domain | Task Type | In Person *P*-value | Remote *P*-value | Remote–In Person *P*-value | In Person PtBis | Remote PtBis | Remote–In Person PtBis |
| 3 | Listening | Listen to a Short Exchange | 0.48 | 0.52 | 0.03 | 0.29 | 0.19 | -0.09 |
| 3 | Listening | Listen to a Classroom Conversation | 0.94 | 0.95 | 0.01 | 0.31 | 0.22 | -0.09 |
| 3 | Listening | Listen to a Classroom Conversation | 0.95 | 0.95 | 0.00 | 0.27 | 0.22 | -0.05 |
| 3 | Listening | Listen to a Classroom Conversation | 0.94 | 0.91 | -0.04 | 0.27 | 0.39 | 0.12 |
| 3 | Listening | Listen to a Story | 0.90 | 0.89 | -0.01 | 0.36 | 0.42 | 0.06 |
| 3 | Listening | Listen to a Story | 0.88 | 0.86 | -0.02 | 0.37 | 0.48 | 0.11 |
| 3 | Listening | Listen to a Story | 0.71 | 0.67 | -0.04 | 0.41 | 0.45 | 0.04 |
| 3 | Listening | Listen to a Story | 0.94 | 0.93 | -0.01 | 0.38 | 0.44 | 0.05 |
| 3 | Listening | Listen to a Story | 0.70 | 0.71 | 0.01 | 0.40 | 0.41 | 0.00 |
| 3 | Listening | Listen to a Story | 0.75 | 0.74 | -0.01 | 0.54 | 0.48 | -0.06 |
| 3 | Listening | Listen to an Oral Presentation | 0.20 | 0.18 | -0.02 | 0.19 | 0.23 | 0.04 |
| 3 | Listening | Listen to an Oral Presentation | 0.63 | 0.61 | -0.02 | 0.40 | 0.37 | -0.04 |
| 3 | Listening | Listen to an Oral Presentation | 0.54 | 0.52 | -0.02 | 0.40 | 0.41 | 0.01 |
| 3 | Listening | Listen to an Oral Presentation | 0.76 | 0.76 | 0.00 | 0.33 | 0.32 | 0.00 |
| 3 | Listening | Listen to an Oral Presentation | 0.93 | 0.97 | 0.04 | 0.45 | 0.17 | -0.27 |
| 3 | Listening | Listen to an Oral Presentation | 0.53 | 0.51 | -0.02 | 0.34 | 0.35 | 0.01 |
| 3 | Listening | Listen to an Oral Presentation | 0.68 | 0.66 | -0.02 | 0.36 | 0.41 | 0.06 |
| 3 | Listening | Listen to an Oral Presentation | 0.60 | 0.60 | 0.00 | 0.46 | 0.33 | -0.13 |
| 3 | Listening | Listen to an Oral Presentation | 0.47 | 0.46 | -0.02 | 0.32 | 0.39 | 0.06 |
| 3 | Listening | Listen to an Oral Presentation | 0.50 | 0.51 | 0.01 | 0.32 | 0.37 | 0.05 |
| 3 | Speaking | Talk About a Scene | 0.97 | 0.99 | 0.02 | 0.39 | 0.30 | -0.09 |
| 3 | Speaking | Talk About a Scene | 0.95 | 0.98 | 0.03 | 0.44 | 0.32 | -0.12 |
| 3 | Speaking | Talk About a Scene | 0.83 | 0.86 | 0.03 | 0.63 | 0.49 | -0.13 |

Table A.1 *(continuation five)*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level | Domain | Task Type | In Person *P*-value | Remote *P*-value | Remote–In Person *P*-value | In Person PtBis | Remote PtBis | Remote–In Person PtBis |
| 3 | Speaking | Talk About a Scene | 0.98 | 0.98 | 0.00 | 0.43 | 0.39 | -0.04 |
| 3 | Speaking | Speech Functions | 0.82 | 0.86 | 0.04 | 0.53 | 0.52 | -0.01 |
| 3 | Speaking | Speech Functions | 0.88 | 0.92 | 0.04 | 0.59 | 0.46 | -0.13 |
| 3 | Speaking | Speech Functions | 0.91 | 0.94 | 0.03 | 0.57 | 0.47 | -0.11 |
| 3 | Speaking | Support an Opinion | 0.86 | 0.89 | 0.03 | 0.72 | 0.64 | -0.08 |
| 3 | Speaking | Support an Opinion | 0.89 | 0.92 | 0.03 | 0.71 | 0.63 | -0.07 |
| 3 | Speaking | Retell a Narrative | 0.84 | 0.88 | 0.03 | 0.74 | 0.66 | -0.08 |
| 3 | Speaking | Summarize an Academic Presentation | 0.71 | 0.75 | 0.04 | 0.77 | 0.75 | -0.02 |
| 3 | Speaking | Summarize an Academic Presentation | 0.70 | 0.74 | 0.04 | 0.74 | 0.73 | -0.01 |
| 6 | Listening | Listen to a Short Exchange | 0.45 | 0.48 | 0.03 | 0.28 | 0.28 | -0.01 |
| 6 | Listening | Listen to a Short Exchange | 0.89 | 0.88 | -0.01 | 0.46 | 0.33 | -0.13 |
| 6 | Listening | Listen to a Short Exchange | 0.89 | 0.87 | -0.02 | 0.44 | 0.50 | 0.06 |
| 6 | Listening | Listen to a Classroom Conversation | 0.89 | 0.88 | -0.02 | 0.30 | 0.43 | 0.13 |
| 6 | Listening | Listen to a Classroom Conversation | 0.92 | 0.87 | -0.05 | 0.30 | 0.46 | 0.16 |
| 6 | Listening | Listen to a Classroom Conversation | 0.95 | 0.93 | -0.02 | 0.22 | 0.31 | 0.09 |
| 6 | Listening | Listen to an Oral Presentation | 0.88 | 0.92 | 0.03 | 0.56 | 0.37 | -0.19 |
| 6 | Listening | Listen to an Oral Presentation | 0.56 | 0.52 | -0.04 | 0.22 | 0.25 | 0.03 |
| 6 | Listening | Listen to an Oral Presentation | 0.61 | 0.59 | -0.03 | 0.36 | 0.43 | 0.07 |
| 6 | Listening | Listen to an Oral Presentation | 0.47 | 0.43 | -0.04 | 0.47 | 0.36 | -0.11 |
| 6 | Listening | Listen to an Oral Presentation | 0.66 | 0.62 | -0.04 | 0.39 | 0.44 | 0.05 |
| 6 | Listening | Listen to an Oral Presentation | 0.66 | 0.68 | 0.01 | 0.40 | 0.37 | -0.03 |

Table A.1 *(continuation six)*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level | Domain | Task Type | In Person *P*-value | Remote *P*-value | Remote–In Person *P*-value | In Person PtBis | Remote PtBis | Remote–In Person PtBis |
| 6 | Listening | Listen to an Oral Presentation | 0.86 | 0.80 | -0.06 | 0.40 | 0.42 | 0.02 |
| 6 | Listening | Listen to an Oral Presentation | 0.90 | 0.87 | -0.03 | 0.47 | 0.35 | -0.13 |
| 6 | Listening | Listen to a Speaker Support an Opinion | 0.63 | 0.61 | -0.02 | 0.30 | 0.36 | 0.06 |
| 6 | Listening | Listen to a Speaker Support an Opinion | 0.71 | 0.66 | -0.05 | 0.51 | 0.59 | 0.07 |
| 6 | Listening | Listen to a Speaker Support an Opinion | 0.82 | 0.71 | -0.12 | 0.27 | 0.49 | 0.22 |
| 6 | Listening | Listen to a Speaker Support an Opinion | 0.91 | 0.86 | -0.05 | 0.29 | 0.34 | 0.05 |
| 6 | Listening | Listen to a Speaker Support an Opinion | 0.90 | 0.85 | -0.05 | 0.37 | 0.39 | 0.03 |
| 6 | Listening | Listen to a Speaker Support an Opinion | 0.54 | 0.52 | -0.02 | 0.42 | 0.47 | 0.05 |
| 6 | Listening | Listen to a Speaker Support an Opinion | 0.68 | 0.68 | 0.00 | 0.49 | 0.41 | -0.07 |
| 6 | Listening | Listen to a Speaker Support an Opinion | 0.31 | 0.27 | -0.04 | 0.26 | 0.33 | 0.07 |
| 6 | Speaking | Talk About a Scene | 1.00 | 0.99 | -0.01 | 0.08 | 0.22 | 0.14 |
| 6 | Speaking | Talk About a Scene | 0.99 | 0.96 | -0.03 | 0.10 | 0.52 | 0.42 |
| 6 | Speaking | Talk About a Scene | 0.95 | 0.92 | -0.04 | 0.35 | 0.61 | 0.26 |
| 6 | Speaking | Talk About a Scene | 0.97 | 0.94 | -0.03 | 0.41 | 0.58 | 0.18 |
| 6 | Speaking | Speech Functions | 0.98 | 0.96 | -0.02 | 0.24 | 0.34 | 0.10 |
| 6 | Speaking | Speech Functions | 0.85 | 0.85 | 0.00 | 0.55 | 0.65 | 0.10 |
| 6 | Speaking | Support an Opinion | 0.81 | 0.81 | 0.00 | 0.73 | 0.83 | 0.09 |
| 6 | Speaking | Support an Opinion | 0.83 | 0.83 | 0.01 | 0.74 | 0.83 | 0.09 |

Table A.1 *(continuation seven)*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level | Domain | Task Type | In Person *P*-value | Remote *P*-value | Remote–In Person *P*-value | In Person PtBis | Remote PtBis | Remote–In Person PtBis |
| 9 | Listening | Listen to a Short Exchange | 0.79 | 0.83 | 0.04 | 0.50 | 0.39 | -0.11 |
| 9 | Listening | Listen to a Short Exchange | 0.92 | 0.82 | -0.10 | 0.37 | 0.42 | 0.06 |
| 9 | Listening | Listen to a Short Exchange | 0.87 | 0.87 | 0.00 | 0.46 | 0.38 | -0.08 |
| 9 | Listening | Listen to a Classroom Conversation | 0.81 | 0.72 | -0.09 | 0.38 | 0.47 | 0.09 |
| 9 | Listening | Listen to a Classroom Conversation | 0.83 | 0.78 | -0.05 | 0.31 | 0.36 | 0.04 |
| 9 | Listening | Listen to a Classroom Conversation | 0.94 | 0.83 | -0.11 | 0.48 | 0.70 | 0.22 |
| 9 | Listening | Listen to an Oral Presentation | 0.68 | 0.65 | -0.03 | 0.41 | 0.57 | 0.16 |
| 9 | Listening | Listen to an Oral Presentation | 0.50 | 0.51 | 0.01 | 0.43 | 0.49 | 0.06 |
| 9 | Listening | Listen to an Oral Presentation | 0.37 | 0.30 | -0.08 | 0.27 | 0.41 | 0.14 |
| 9 | Listening | Listen to an Oral Presentation | 0.79 | 0.72 | -0.07 | 0.36 | 0.47 | 0.10 |
| 9 | Listening | Listen to an Oral Presentation | 0.68 | 0.61 | -0.06 | 0.47 | 0.38 | -0.09 |
| 9 | Listening | Listen to an Oral Presentation | 0.55 | 0.57 | 0.02 | 0.31 | 0.32 | 0.02 |
| 9 | Listening | Listen to an Oral Presentation | 0.55 | 0.48 | -0.06 | 0.22 | 0.32 | 0.10 |
| 9 | Listening | Listen to an Oral Presentation | 0.60 | 0.47 | -0.13 | 0.48 | 0.57 | 0.09 |
| 9 | Listening | Listen to a Speaker Support an Opinion | 0.94 | 0.94 | 0.00 | 0.41 | 0.26 | -0.15 |
| 9 | Listening | Listen to a Speaker Support an Opinion | 0.24 | 0.27 | 0.04 | 0.37 | 0.32 | -0.04 |
| 9 | Listening | Listen to a Speaker Support an Opinion | 0.61 | 0.54 | -0.07 | 0.36 | 0.21 | -0.15 |
| 9 | Listening | Listen to a Speaker Support an Opinion | 0.58 | 0.57 | -0.02 | 0.53 | 0.44 | -0.09 |
| 9 | Listening | Listen to a Speaker Support an Opinion | 0.73 | 0.75 | 0.02 | 0.47 | 0.31 | -0.15 |
| 9 | Listening | Listen to a Speaker Support an Opinion | 0.68 | 0.61 | -0.06 | 0.50 | 0.50 | 0.00 |
| 9 | Listening | Listen to a Speaker Support an Opinion | 0.83 | 0.83 | 0.00 | 0.37 | 0.34 | -0.03 |
| 9 | Listening | Listen to a Speaker Support an Opinion | 0.91 | 0.83 | -0.08 | 0.39 | 0.61 | 0.22 |

Table A.1 *(continuation eight)*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level | Domain | Task Type | In Person *P*-value | Remote *P*-value | Remote–In Person *P*-value | In Person PtBis | Remote PtBis | Remote–In Person PtBis |
| 9 | Speaking | Talk About a Scene | 0.93 | 0.98 | 0.05 | 0.72 | 0.36 | -0.36 |
| 9 | Speaking | Talk About a Scene | 0.92 | 0.95 | 0.03 | 0.60 | 0.47 | -0.13 |
| 9 | Speaking | Talk About a Scene | 0.87 | 0.92 | 0.05 | 0.72 | 0.52 | -0.20 |
| 9 | Speaking | Talk About a Scene | 0.89 | 0.95 | 0.06 | 0.78 | 0.34 | -0.44 |
| 9 | Speaking | Speech Functions | 0.84 | 0.89 | 0.05 | 0.75 | 0.60 | -0.16 |
| 9 | Speaking | Speech Functions | 0.69 | 0.77 | 0.08 | 0.70 | 0.69 | -0.01 |
| 9 | Speaking | Support an Opinion | 0.79 | 0.84 | 0.05 | 0.88 | 0.78 | -0.10 |
| 9 | Speaking | Support an Opinion | 0.78 | 0.82 | 0.04 | 0.86 | 0.82 | -0.04 |
| 9 | Speaking | Present and Discuss Information | 0.78 | 0.81 | 0.04 | 0.85 | 0.78 | -0.06 |
| 9 | Speaking | Present and Discuss Information | 0.72 | 0.73 | 0.01 | 0.80 | 0.76 | -0.04 |
| 9 | Speaking | Summarize an Academic Presentation | 0.68 | 0.70 | 0.02 | 0.85 | 0.87 | 0.03 |
| 9 | Speaking | Summarize an Academic Presentation | 0.57 | 0.61 | 0.04 | 0.78 | 0.80 | 0.02 |
| 11 | Listening | Listen to a Short Exchange | 0.66 | 0.64 | -0.02 | 0.12 | 0.08 | -0.04 |
| 11 | Listening | Listen to a Short Exchange | 0.70 | 0.75 | 0.05 | 0.43 | 0.59 | 0.16 |
| 11 | Listening | Listen to a Short Exchange | 0.86 | 0.70 | -0.15 | 0.43 | 0.54 | 0.11 |
| 11 | Listening | Listen to a Classroom Conversation | 0.53 | 0.68 | 0.15 | 0.26 | 0.47 | 0.21 |
| 11 | Listening | Listen to a Classroom Conversation | 0.50 | 0.49 | -0.01 | 0.47 | 0.38 | -0.09 |
| 11 | Listening | Listen to a Classroom Conversation | 0.74 | 0.78 | 0.04 | 0.48 | 0.52 | 0.05 |
| 11 | Listening | Listen to an Oral Presentation | 0.74 | 0.64 | -0.10 | 0.45 | 0.55 | 0.10 |
| 11 | Listening | Listen to an Oral Presentation | 0.71 | 0.72 | 0.01 | 0.42 | 0.26 | -0.16 |
| 11 | Listening | Listen to an Oral Presentation | 0.55 | 0.70 | 0.14 | 0.71 | 0.50 | -0.21 |
| 11 | Listening | Listen to an Oral Presentation | 0.48 | 0.54 | 0.05 | 0.60 | 0.58 | -0.02 |
| 11 | Listening | Listen to an Oral Presentation | 0.63 | 0.54 | -0.09 | 0.34 | 0.28 | -0.06 |
| 11 | Listening | Listen to an Oral Presentation | 0.61 | 0.49 | -0.12 | 0.53 | 0.21 | -0.32 |
| 11 | Listening | Listen to an Oral Presentation | 0.63 | 0.69 | 0.05 | 0.39 | 0.43 | 0.03 |
| 11 | Listening | Listen to an Oral Presentation | 0.60 | 0.55 | -0.05 | 0.34 | 0.51 | 0.16 |

Table A.1 *(continuation nine)*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level | Domain | Task Type | In Person *P*-value | Remote *P*-value | Remote–In Person *P*-value | In Person PtBis | Remote PtBis | Remote–In Person PtBis |
| 11 | Listening | Listen to a Speaker Support an Opinion | 0.93 | 0.87 | -0.06 | 0.44 | 0.60 | 0.16 |
| 11 | Listening | Listen to a Speaker Support an Opinion | 0.70 | 0.78 | 0.08 | 0.66 | 0.50 | -0.16 |
| 11 | Listening | Listen to a Speaker Support an Opinion | 0.74 | 0.89 | 0.15 | 0.29 | 0.09 | -0.20 |
| 11 | Listening | Listen to a Speaker Support an Opinion | 0.74 | 0.86 | 0.13 | 0.58 | 0.15 | -0.43 |
| 11 | Listening | Listen to a Speaker Support an Opinion | 0.76 | 0.91 | 0.16 | 0.50 | 0.18 | -0.32 |
| 11 | Listening | Listen to a Speaker Support an Opinion | 0.86 | 0.90 | 0.04 | 0.46 | 0.24 | -0.21 |
| 11 | Listening | Listen to a Speaker Support an Opinion | 0.74 | 0.66 | -0.07 | 0.29 | 0.52 | 0.23 |
| 11 | Listening | Listen to a Speaker Support an Opinion | 0.71 | 0.71 | 0.00 | 0.65 | 0.58 | -0.07 |
| 11 | Speaking | Talk About a Scene | 1.00 | 0.90 | -0.09 | -0.06 | 0.69 | 0.74 |
| 11 | Speaking | Talk About a Scene | 1.00 | 0.91 | -0.09 | N/A | 0.70 | N/A |
| 11 | Speaking | Talk About a Scene | 0.97 | 0.95 | -0.02 | 0.15 | 0.67 | 0.53 |
| 11 | Speaking | Talk About a Scene | 0.95 | 0.88 | -0.07 | 0.37 | 0.69 | 0.33 |
| 11 | Speaking | Speech Functions | 0.96 | 0.80 | -0.16 | 0.28 | 0.87 | 0.59 |
| 11 | Speaking | Speech Functions | 0.93 | 0.84 | -0.09 | 0.71 | 0.80 | 0.09 |
| 11 | Speaking | Support an Opinion | 0.85 | 0.84 | -0.02 | 0.80 | 0.85 | 0.04 |
| 11 | Speaking | Support an Opinion | 0.86 | 0.82 | -0.04 | 0.41 | 0.89 | 0.48 |
| 11 | Speaking | Present and Discuss Information | 0.76 | 0.81 | 0.05 | 0.75 | 0.87 | 0.12 |
| 11 | Speaking | Present and Discuss Information | 0.72 | 0.75 | 0.03 | 0.54 | 0.82 | 0.29 |
| 11 | Speaking | Summarize an Academic Presentation | 0.68 | 0.63 | -0.05 | 0.82 | 0.87 | 0.05 |
| 11 | Speaking | Summarize an Academic Presentation | 0.72 | 0.73 | 0.01 | 0.84 | 0.60 | -0.24 |

Table A.2 Written Language Classical Item Statistics

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level | Domain | Task Type | In Person *P*-value | Remote *P*-value | Remote–In Person *P*-value | In Person PtBis | Remote PtBis | Remote–In Person PtBis |
| Kindergarten | Reading | Read-Along Word with Scaffolding | 0.83 | 0.89 | 0.06 | 0.75 | 0.70 | -0.04 |
| Kindergarten | Reading | Read-Along Word with Scaffolding | 0.88 | 0.94 | 0.06 | 0.58 | 0.46 | -0.12 |
| Kindergarten | Reading | Read-Along Word with Scaffolding | 0.80 | 0.84 | 0.04 | 0.69 | 0.57 | -0.13 |
| Kindergarten | Reading | Read-Along Word with Scaffolding | 0.95 | 0.99 | 0.04 | 0.42 | 0.00 | -0.41 |
| Kindergarten | Reading | Read-Along Story with Scaffolding | 0.81 | 0.89 | 0.08 | 0.74 | 0.58 | -0.16 |
| Kindergarten | Reading | Read-Along Story with Scaffolding | 0.66 | 0.71 | 0.05 | 0.63 | 0.63 | 0.00 |
| Kindergarten | Reading | Read-Along Story with Scaffolding | 0.75 | 0.79 | 0.04 | 0.58 | 0.61 | 0.02 |
| Kindergarten | Reading | Read-Along Story with Scaffolding | 0.75 | 0.82 | 0.07 | 0.58 | 0.64 | 0.06 |
| Kindergarten | Reading | Read-Along Information | 0.83 | 0.89 | 0.06 | 0.62 | 0.64 | 0.02 |
| Kindergarten | Reading | Read-Along Information | 0.87 | 0.85 | -0.01 | 0.42 | 0.52 | 0.10 |
| Kindergarten | Reading | Read-Along Information | 0.91 | 0.91 | 0.01 | 0.41 | 0.59 | 0.18 |
| Kindergarten | Reading | Read-Along Information | 0.93 | 0.93 | 0.01 | 0.41 | 0.49 | 0.09 |
| Kindergarten | Reading | Read-Along Information | 0.78 | 0.92 | 0.14 | 0.69 | 0.59 | -0.10 |
| Kindergarten | Reading | Read-Along Information | 0.87 | 0.89 | 0.02 | 0.44 | 0.48 | 0.04 |
| Kindergarten | Writing | Label a Picture—Word with Scaffolding | 0.87 | 0.94 | 0.07 | 0.63 | 0.61 | -0.02 |
| Kindergarten | Writing | Label a Picture—Word with Scaffolding | 0.94 | 0.97 | 0.03 | 0.49 | 0.64 | 0.15 |
| Kindergarten | Writing | Label a Picture—Word with Scaffolding | 0.83 | 0.93 | 0.10 | 0.86 | 0.81 | -0.05 |
| Kindergarten | Writing | Label a Picture—Word with Scaffolding | 0.85 | 0.93 | 0.08 | 0.87 | 0.79 | -0.08 |
| Kindergarten | Writing | Write a Story Together with Scaffolding | 0.76 | 0.84 | 0.08 | 0.73 | 0.59 | -0.14 |
| Kindergarten | Writing | Write a Story Together with Scaffolding | 0.84 | 0.96 | 0.12 | 0.71 | 0.75 | 0.04 |
| Kindergarten | Writing | Write a Story Together with Scaffolding | 0.83 | 0.95 | 0.12 | 0.92 | 0.68 | -0.23 |
| Kindergarten | Writing | Write a Story Together with Scaffolding | 0.84 | 0.94 | 0.11 | 0.91 | 0.83 | -0.08 |

Table A.2 *(continuation one)*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level | Domain | Task Type | In Person *P*-value | Remote *P*-value | Remote–In Person *P*-value | In Person PtBis | Remote PtBis | Remote–In Person PtBis |
| 1 | Reading | Read and Choose a Word | 0.92 | 0.89 | -0.03 | 0.51 | 0.63 | 0.12 |
| 1 | Reading | Read and Choose a Sentence | 0.91 | 0.90 | 0.00 | 0.49 | 0.47 | -0.02 |
| 1 | Reading | Read and Choose a Sentence | 0.94 | 0.92 | -0.02 | 0.50 | 0.56 | 0.06 |
| 1 | Reading | Read and Choose a Sentence | 0.88 | 0.85 | -0.03 | 0.58 | 0.61 | 0.03 |
| 1 | Reading | Read and Choose a Sentence | 0.82 | 0.84 | 0.02 | 0.40 | 0.47 | 0.07 |
| 1 | Reading | Read a Short Informational Passage | 0.87 | 0.88 | 0.01 | 0.56 | 0.65 | 0.09 |
| 1 | Reading | Read a Short Informational Passage | 0.74 | 0.77 | 0.03 | 0.55 | 0.68 | 0.14 |
| 1 | Reading | Read a Short Informational Passage | 0.76 | 0.82 | 0.05 | 0.60 | 0.65 | 0.05 |
| 1 | Reading | Read a Short Informational Passage | 0.87 | 0.84 | -0.03 | 0.60 | 0.68 | 0.07 |
| 1 | Reading | Read a Short Informational Passage | 0.68 | 0.73 | 0.05 | 0.63 | 0.57 | -0.07 |
| 1 | Reading | Read a Short Informational Passage | 0.70 | 0.71 | 0.02 | 0.65 | 0.60 | -0.05 |
| 1 | Reading | Read a Literary Passage | 0.89 | 0.88 | -0.01 | 0.64 | 0.68 | 0.04 |
| 1 | Reading | Read a Literary Passage | 0.59 | 0.63 | 0.04 | 0.56 | 0.49 | -0.06 |
| 1 | Reading | Read a Literary Passage | 0.79 | 0.81 | 0.02 | 0.61 | 0.59 | -0.02 |
| 1 | Reading | Read an Informational Passage | 0.74 | 0.72 | -0.01 | 0.60 | 0.67 | 0.07 |
| 1 | Reading | Read an Informational Passage | 0.79 | 0.87 | 0.08 | 0.58 | 0.54 | -0.03 |
| 1 | Reading | Read an Informational Passage | 0.61 | 0.63 | 0.02 | 0.53 | 0.50 | -0.03 |
| 1 | Reading | Read an Informational Passage | 0.67 | 0.69 | 0.02 | 0.63 | 0.68 | 0.05 |
| 1 | Reading | Read an Informational Passage | 0.53 | 0.56 | 0.02 | 0.59 | 0.38 | -0.20 |
| 1 | Reading | Read an Informational Passage | 0.70 | 0.66 | -0.03 | 0.67 | 0.70 | 0.03 |

Table A.2 *(continuation two)*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level | Domain | Task Type | In Person *P*-value | Remote *P*-value | Remote–In Person *P*-value | In Person PtBis | Remote PtBis | Remote–In Person PtBis |
| 1 | Writing | Write a Story Together with Scaffolding | 0.95 | 0.97 | 0.03 | 0.32 | 0.41 | 0.09 |
| 1 | Writing | Write a Story Together with Scaffolding | 0.90 | 0.86 | -0.04 | 0.48 | 0.55 | 0.07 |
| 1 | Writing | Write a Story Together with Scaffolding | 0.71 | 0.76 | 0.05 | 0.74 | 0.79 | 0.05 |
| 1 | Writing | Write an Informational Text Together | 0.74 | 0.78 | 0.04 | 0.70 | 0.65 | -0.05 |
| 1 | Writing | Write an Informational Text Together | 0.72 | 0.76 | 0.03 | 0.78 | 0.75 | -0.03 |
| 1 | Writing | Describe a Picture | 0.70 | 0.76 | 0.07 | 0.83 | 0.76 | -0.07 |
| 1 | Writing | Write About an Experience | 0.66 | 0.69 | 0.03 | 0.79 | 0.77 | -0.02 |
| 2 | Reading | Read and Choose a Sentence | 0.97 | 0.97 | 0.00 | 0.29 | 0.34 | 0.05 |
| 2 | Reading | Read and Choose a Sentence | 0.93 | 0.90 | -0.03 | 0.28 | 0.18 | -0.10 |
| 2 | Reading | Read and Choose a Sentence | 0.99 | 0.99 | 0.00 | 0.13 | 0.12 | 0.00 |
| 2 | Reading | Read and Choose a Sentence | 1.00 | 1.00 | 0.00 | 0.19 | N/A | N/A |
| 2 | Reading | Read a Short Informational Passage | 0.81 | 0.80 | 0.00 | 0.38 | 0.49 | 0.11 |
| 2 | Reading | Read a Short Informational Passage | 0.96 | 0.96 | 0.00 | 0.20 | 0.30 | 0.11 |
| 2 | Reading | Read a Short Informational Passage | 0.88 | 0.86 | -0.02 | 0.32 | 0.49 | 0.17 |
| 2 | Reading | Read a Short Informational Passage | 0.93 | 0.95 | 0.02 | 0.47 | 0.35 | -0.12 |
| 2 | Reading | Read a Short Informational Passage | 0.96 | 0.97 | 0.01 | 0.32 | 0.33 | 0.01 |
| 2 | Reading | Read a Short Informational Passage | 0.40 | 0.38 | -0.02 | 0.33 | 0.37 | 0.03 |
| 2 | Reading | Read a Literary Passage | 0.97 | 0.98 | 0.01 | 0.53 | 0.34 | -0.18 |
| 2 | Reading | Read a Literary Passage | 0.83 | 0.84 | 0.01 | 0.43 | 0.42 | -0.01 |
| 2 | Reading | Read a Literary Passage | 0.95 | 0.95 | 0.00 | 0.51 | 0.47 | -0.04 |
| 2 | Reading | Read a Literary Passage | 0.90 | 0.88 | -0.02 | 0.47 | 0.32 | -0.15 |
| 2 | Reading | Read a Literary Passage | 0.90 | 0.93 | 0.03 | 0.40 | 0.52 | 0.11 |
| 2 | Reading | Read a Literary Passage | 0.83 | 0.80 | -0.03 | 0.48 | 0.37 | -0.11 |

Table A.2 *(continuation three)*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level | Domain | Task Type | In Person *P*-value | Remote *P*-value | Remote–In Person *P*-value | In Person PtBis | Remote PtBis | Remote–In Person PtBis |
| 2 | Reading | Read a Literary Passage | 0.88 | 0.91 | 0.03 | 0.50 | 0.54 | 0.04 |
| 2 | Reading | Read a Literary Passage | 0.55 | 0.52 | -0.03 | 0.31 | 0.31 | 0.00 |
| 2 | Reading | Read an Informational Passage | 0.89 | 0.94 | 0.05 | 0.46 | 0.20 | -0.26 |
| 2 | Reading | Read an Informational Passage | 0.84 | 0.88 | 0.04 | 0.63 | 0.58 | -0.04 |
| 2 | Reading | Read an Informational Passage | 0.87 | 0.89 | 0.01 | 0.55 | 0.53 | -0.02 |
| 2 | Reading | Read an Informational Passage | 0.90 | 0.88 | -0.02 | 0.58 | 0.59 | 0.01 |
| 2 | Reading | Read an Informational Passage | 0.77 | 0.83 | 0.06 | 0.52 | 0.54 | 0.02 |
| 2 | Reading | Read an Informational Passage | 0.81 | 0.78 | -0.03 | 0.53 | 0.54 | 0.01 |
| 2 | Reading | Read an Informational Passage | 0.82 | 0.86 | 0.04 | 0.49 | 0.54 | 0.05 |
| 2 | Reading | Read an Informational Passage | 0.88 | 0.81 | -0.06 | 0.60 | 0.55 | -0.05 |
| 2 | Writing | Write a Story Together with Scaffolding | 0.84 | 0.83 | -0.01 | 0.62 | 0.55 | -0.07 |
| 2 | Writing | Write a Story Together with Scaffolding | 0.86 | 0.90 | 0.04 | 0.63 | 0.51 | -0.11 |
| 2 | Writing | Write an Informational Text Together | 0.83 | 0.83 | 0.00 | 0.63 | 0.67 | 0.04 |
| 2 | Writing | Write an Informational Text Together | 0.85 | 0.86 | 0.01 | 0.68 | 0.74 | 0.06 |
| 2 | Writing | Describe a Picture | 0.90 | 0.88 | -0.02 | 0.70 | 0.63 | -0.07 |
| 2 | Writing | Write About an Experience | 0.75 | 0.70 | -0.04 | 0.73 | 0.73 | 0.00 |
| 3 | Reading | Read and Choose a Sentence | 0.90 | 0.87 | -0.03 | 0.38 | 0.41 | 0.03 |
| 3 | Reading | Read and Choose a Sentence | 0.92 | 0.92 | 0.00 | 0.41 | 0.39 | -0.01 |
| 3 | Reading | Read a Short Informational Passage | 0.63 | 0.64 | 0.00 | 0.45 | 0.36 | -0.09 |
| 3 | Reading | Read a Short Informational Passage | 0.53 | 0.50 | -0.03 | 0.42 | 0.43 | 0.01 |
| 3 | Reading | Read a Short Informational Passage | 0.52 | 0.53 | 0.01 | 0.51 | 0.44 | -0.07 |
| 3 | Reading | Read a Short Informational Passage | 0.53 | 0.52 | -0.02 | 0.53 | 0.51 | -0.03 |
| 3 | Reading | Read a Short Informational Passage | 0.71 | 0.66 | -0.04 | 0.53 | 0.53 | 0.00 |
| 3 | Reading | Read a Short Informational Passage | 0.54 | 0.50 | -0.04 | 0.46 | 0.51 | 0.05 |

Table A.2 *(continuation four)*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level | Domain | Task Type | In Person *P*-value | Remote *P*-value | Remote–In Person *P*-value | In Person PtBis | Remote PtBis | Remote–In Person PtBis |
| 3 | Reading | Read a Student Essay | 0.32 | 0.31 | -0.01 | 0.29 | 0.29 | 0.00 |
| 3 | Reading | Read a Student Essay | 0.34 | 0.36 | 0.02 | 0.38 | 0.35 | -0.03 |
| 3 | Reading | Read a Student Essay | 0.36 | 0.35 | -0.01 | 0.28 | 0.24 | -0.04 |
| 3 | Reading | Read a Student Essay | 0.27 | 0.29 | 0.01 | 0.27 | 0.24 | -0.03 |
| 3 | Reading | Read a Student Essay | 0.52 | 0.50 | -0.02 | 0.36 | 0.42 | 0.05 |
| 3 | Reading | Read a Student Essay | 0.58 | 0.56 | -0.01 | 0.50 | 0.47 | -0.02 |
| 3 | Reading | Read a Literary Passage | 0.59 | 0.59 | -0.01 | 0.59 | 0.51 | -0.09 |
| 3 | Reading | Read a Literary Passage | 0.70 | 0.72 | 0.02 | 0.37 | 0.35 | -0.02 |
| 3 | Reading | Read a Literary Passage | 0.70 | 0.72 | 0.02 | 0.46 | 0.38 | -0.08 |
| 3 | Reading | Read a Literary Passage | 0.39 | 0.39 | 0.00 | 0.36 | 0.41 | 0.05 |
| 3 | Reading | Read a Literary Passage | 0.50 | 0.50 | 0.00 | 0.46 | 0.39 | -0.07 |
| 3 | Reading | Read a Literary Passage | 0.68 | 0.67 | -0.01 | 0.58 | 0.57 | -0.01 |
| 3 | Reading | Read an Informational Passage | 0.64 | 0.60 | -0.04 | 0.58 | 0.63 | 0.05 |
| 3 | Reading | Read an Informational Passage | 0.43 | 0.46 | 0.03 | 0.27 | 0.23 | -0.04 |
| 3 | Reading | Read an Informational Passage | 0.54 | 0.52 | -0.02 | 0.45 | 0.51 | 0.07 |
| 3 | Reading | Read an Informational Passage | 0.30 | 0.34 | 0.04 | 0.33 | 0.26 | -0.08 |
| 3 | Reading | Read an Informational Passage | 0.32 | 0.37 | 0.05 | 0.27 | 0.27 | 0.00 |
| 3 | Reading | Read an Informational Passage | 0.39 | 0.39 | 0.00 | 0.18 | 0.21 | 0.03 |
| 3 | Writing | Describe a Picture | 0.75 | 0.80 | 0.05 | 0.68 | 0.55 | -0.13 |
| 3 | Writing | Describe a Picture | 0.74 | 0.77 | 0.03 | 0.68 | 0.63 | -0.06 |
| 3 | Writing | Write About an Experience | 0.61 | 0.66 | 0.05 | 0.85 | 0.78 | -0.07 |
| 3 | Writing | Write About Academic Information | 0.63 | 0.64 | 0.01 | 0.67 | 0.68 | 0.01 |
| 3 | Writing | Write About Academic Information | 0.65 | 0.66 | 0.01 | 0.80 | 0.76 | -0.04 |
| 3 | Writing | Justify an Opinion | 0.56 | 0.59 | 0.04 | 0.84 | 0.81 | -0.03 |
| 6 | Reading | Read a Short Informational Passage | 0.63 | 0.62 | -0.01 | 0.35 | 0.37 | 0.02 |
| 6 | Reading | Read a Short Informational Passage | 0.47 | 0.54 | 0.07 | 0.36 | 0.22 | -0.14 |
| 6 | Reading | Read a Short Informational Passage | 0.38 | 0.37 | -0.01 | 0.46 | 0.38 | -0.07 |

Table A.2 *(continuation five)*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level | Domain | Task Type | In Person *P*-value | Remote *P*-value | Remote–In Person *P*-value | In Person PtBis | Remote PtBis | Remote–In Person PtBis |
| 6 | Reading | Read a Short Informational Passage | 0.57 | 0.53 | -0.04 | 0.24 | 0.40 | 0.16 |
| 6 | Reading | Read a Short Informational Passage | 0.35 | 0.29 | -0.06 | 0.22 | 0.23 | 0.02 |
| 6 | Reading | Read a Short Informational Passage | 0.64 | 0.62 | -0.02 | 0.50 | 0.42 | -0.08 |
| 6 | Reading | Read a Student Essay | 0.54 | 0.58 | 0.04 | 0.57 | 0.46 | -0.11 |
| 6 | Reading | Read a Student Essay | 0.30 | 0.34 | 0.04 | 0.30 | 0.16 | -0.14 |
| 6 | Reading | Read a Student Essay | 0.56 | 0.52 | -0.04 | 0.20 | 0.33 | 0.13 |
| 6 | Reading | Read a Student Essay | 0.73 | 0.66 | -0.07 | 0.59 | 0.57 | -0.03 |
| 6 | Reading | Read a Student Essay | 0.62 | 0.67 | 0.05 | 0.52 | 0.38 | -0.14 |
| 6 | Reading | Read a Student Essay | 0.60 | 0.55 | -0.04 | 0.49 | 0.54 | 0.05 |
| 6 | Reading | Read a Student Essay | 0.31 | 0.30 | -0.01 | 0.32 | 0.33 | 0.02 |
| 6 | Reading | Read a Student Essay | 0.42 | 0.47 | 0.05 | 0.33 | 0.23 | -0.10 |
| 6 | Reading | Read a Literary Passage | 0.31 | 0.26 | -0.05 | 0.06 | 0.30 | 0.24 |
| 6 | Reading | Read a Literary Passage | 0.27 | 0.28 | 0.01 | 0.42 | 0.50 | 0.08 |
| 6 | Reading | Read a Literary Passage | 0.37 | 0.38 | 0.01 | 0.40 | 0.42 | 0.02 |
| 6 | Reading | Read a Literary Passage | 0.54 | 0.47 | -0.07 | 0.54 | 0.61 | 0.07 |
| 6 | Reading | Read a Literary Passage | 0.34 | 0.33 | -0.01 | 0.41 | 0.37 | -0.04 |
| 6 | Reading | Read a Literary Passage | 0.35 | 0.31 | -0.04 | 0.39 | 0.45 | 0.06 |
| 6 | Reading | Read an Informational Passage | 0.36 | 0.33 | -0.02 | 0.31 | 0.33 | 0.01 |
| 6 | Reading | Read an Informational Passage | 0.23 | 0.25 | 0.02 | 0.33 | 0.24 | -0.09 |
| 6 | Reading | Read an Informational Passage | 0.41 | 0.36 | -0.05 | 0.43 | 0.51 | 0.08 |
| 6 | Reading | Read an Informational Passage | 0.44 | 0.44 | 0.00 | 0.42 | 0.37 | -0.05 |
| 6 | Reading | Read an Informational Passage | 0.45 | 0.42 | -0.03 | 0.53 | 0.56 | 0.02 |
| 6 | Reading | Read an Informational Passage | 0.37 | 0.35 | -0.02 | 0.24 | 0.24 | 0.00 |

Table A.2 *(continuation six)*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level | Domain | Task Type | In Person *P*-value | Remote *P*-value | Remote–In Person *P*-value | In Person PtBis | Remote PtBis | Remote–In Person PtBis |
| 6 | Writing | Describe a Picture | 0.70 | 0.74 | 0.04 | 0.74 | 0.57 | -0.16 |
| 6 | Writing | Describe a Picture | 0.81 | 0.83 | 0.03 | 0.61 | 0.46 | -0.15 |
| 6 | Writing | Write About an Experience | 0.68 | 0.70 | 0.02 | 0.84 | 0.76 | -0.08 |
| 6 | Writing | Write About Academic Information | 0.83 | 0.79 | -0.03 | 0.66 | 0.65 | 0.00 |
| 6 | Writing | Write About Academic Information | 0.68 | 0.68 | 0.00 | 0.76 | 0.73 | -0.03 |
| 6 | Writing | Justify an Opinion | 0.59 | 0.61 | 0.02 | 0.87 | 0.83 | -0.04 |
| 9 | Reading | Read a Short Informational Passage | 0.66 | 0.66 | 0.00 | 0.43 | 0.39 | -0.04 |
| 9 | Reading | Read a Short Informational Passage | 0.61 | 0.57 | -0.04 | 0.53 | 0.49 | -0.04 |
| 9 | Reading | Read a Short Informational Passage | 0.69 | 0.69 | 0.01 | 0.57 | 0.51 | -0.06 |
| 9 | Reading | Read a Short Informational Passage | 0.77 | 0.74 | -0.03 | 0.43 | 0.50 | 0.07 |
| 9 | Reading | Read a Short Informational Passage | 0.39 | 0.38 | -0.01 | 0.28 | 0.33 | 0.06 |
| 9 | Reading | Read a Short Informational Passage | 0.50 | 0.49 | -0.01 | 0.45 | 0.52 | 0.06 |
| 9 | Reading | Read a Student Essay | 0.64 | 0.63 | -0.01 | 0.42 | 0.45 | 0.03 |
| 9 | Reading | Read a Student Essay | 0.59 | 0.59 | 0.00 | 0.35 | 0.40 | 0.05 |
| 9 | Reading | Read a Student Essay | 0.43 | 0.36 | -0.07 | 0.50 | 0.44 | -0.06 |
| 9 | Reading | Read a Student Essay | 0.49 | 0.46 | -0.03 | 0.36 | 0.43 | 0.07 |
| 9 | Reading | Read a Student Essay | 0.54 | 0.47 | -0.06 | 0.39 | 0.44 | 0.05 |
| 9 | Reading | Read a Student Essay | 0.62 | 0.56 | -0.06 | 0.38 | 0.38 | -0.01 |
| 9 | Reading | Read a Student Essay | 0.73 | 0.74 | 0.01 | 0.35 | 0.24 | -0.11 |
| 9 | Reading | Read a Student Essay | 0.33 | 0.42 | 0.08 | 0.25 | 0.31 | 0.07 |
| 9 | Reading | Read a Literary Passage | 0.76 | 0.76 | 0.00 | 0.34 | 0.42 | 0.07 |
| 9 | Reading | Read a Literary Passage | 0.50 | 0.52 | 0.02 | 0.60 | 0.52 | -0.08 |
| 9 | Reading | Read a Literary Passage | 0.39 | 0.38 | -0.01 | 0.48 | 0.35 | -0.13 |
| 9 | Reading | Read a Literary Passage | 0.83 | 0.81 | -0.02 | 0.50 | 0.47 | -0.03 |
| 9 | Reading | Read a Literary Passage | 0.82 | 0.81 | -0.01 | 0.53 | 0.56 | 0.03 |
| 9 | Reading | Read a Literary Passage | 0.56 | 0.57 | 0.01 | 0.57 | 0.57 | 0.01 |

Table A.2 *(continuation seven)*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level | Domain | Task Type | In Person *P*-value | Remote *P*-value | Remote–In Person *P*-Value | In Person PtBis | Remote PtBis | Remote–In Person PtBis |
| 9 | Reading | Read an Informational Passage | 0.55 | 0.53 | -0.02 | 0.25 | 0.24 | -0.01 |
| 9 | Reading | Read an Informational Passage | 0.64 | 0.68 | 0.03 | 0.35 | 0.50 | 0.15 |
| 9 | Reading | Read an Informational Passage | 0.58 | 0.58 | 0.00 | 0.36 | 0.40 | 0.04 |
| 9 | Reading | Read an Informational Passage | 0.47 | 0.48 | 0.01 | 0.34 | 0.37 | 0.03 |
| 9 | Reading | Read an Informational Passage | 0.51 | 0.48 | -0.03 | 0.39 | 0.55 | 0.16 |
| 9 | Reading | Read an Informational Passage | 0.71 | 0.66 | -0.05 | 0.52 | 0.61 | 0.09 |
| 9 | Writing | Describe a Picture | 0.76 | 0.78 | 0.02 | 0.56 | 0.48 | -0.09 |
| 9 | Writing | Describe a Picture | 0.83 | 0.85 | 0.02 | 0.54 | 0.56 | 0.01 |
| 9 | Writing | Write About an Experience | 0.76 | 0.77 | 0.01 | 0.80 | 0.73 | -0.06 |
| 9 | Writing | Write About Academic Information | 0.80 | 0.81 | 0.01 | 0.56 | 0.58 | 0.02 |
| 9 | Writing | Write About Academic Information | 0.73 | 0.73 | 0.01 | 0.76 | 0.74 | -0.02 |
| 9 | Writing | Justify an Opinion | 0.66 | 0.68 | 0.02 | 0.83 | 0.81 | -0.03 |
| 11 | Reading | Read a Short Informational Passage | 0.89 | 0.85 | -0.04 | 0.36 | 0.39 | 0.03 |
| 11 | Reading | Read a Short Informational Passage | 0.82 | 0.85 | 0.03 | 0.58 | 0.48 | -0.10 |
| 11 | Reading | Read a Short Informational Passage | 0.41 | 0.43 | 0.02 | 0.47 | 0.42 | -0.05 |
| 11 | Reading | Read a Short Informational Passage | 0.48 | 0.54 | 0.05 | 0.20 | 0.29 | 0.08 |
| 11 | Reading | Read a Short Informational Passage | 0.20 | 0.21 | 0.01 | 0.27 | 0.32 | 0.04 |
| 11 | Reading | Read a Short Informational Passage | 0.43 | 0.44 | 0.01 | 0.31 | 0.47 | 0.16 |
| 11 | Reading | Read a Student Essay | 0.46 | 0.46 | 0.00 | 0.49 | 0.60 | 0.11 |
| 11 | Reading | Read a Student Essay | 0.59 | 0.66 | 0.08 | 0.38 | 0.40 | 0.02 |
| 11 | Reading | Read a Student Essay | 0.60 | 0.46 | -0.14 | 0.29 | 0.44 | 0.15 |
| 11 | Reading | Read a Student Essay | 0.45 | 0.42 | -0.03 | 0.67 | 0.45 | -0.21 |
| 11 | Reading | Read a Student Essay | 0.67 | 0.78 | 0.11 | 0.62 | 0.49 | -0.13 |
| 11 | Reading | Read a Student Essay | 0.69 | 0.76 | 0.08 | 0.58 | 0.53 | -0.05 |
| 11 | Reading | Read a Student Essay | 0.77 | 0.83 | 0.07 | 0.49 | 0.24 | -0.25 |
| 11 | Reading | Read a Student Essay | 0.32 | 0.34 | 0.01 | 0.29 | 0.36 | 0.07 |

Table A.2 *(continuation eight)*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level | Domain | Task Type | In Person *P*-value | Remote *P*-value | Remote–In Person *P*-value | In Person PtBis | Remote PtBis | Remote–In Person PtBis |
| 11 | Reading | Read a Literary Passage | 0.56 | 0.65 | 0.09 | 0.67 | 0.58 | -0.09 |
| 11 | Reading | Read a Literary Passage | 0.56 | 0.55 | -0.01 | 0.46 | 0.25 | -0.20 |
| 11 | Reading | Read a Literary Passage | 0.64 | 0.65 | 0.00 | 0.56 | 0.51 | -0.05 |
| 11 | Reading | Read a Literary Passage | 0.54 | 0.49 | -0.06 | 0.22 | 0.47 | 0.25 |
| 11 | Reading | Read a Literary Passage | 0.55 | 0.49 | -0.05 | 0.51 | 0.58 | 0.07 |
| 11 | Reading | Read a Literary Passage | 0.47 | 0.49 | 0.03 | 0.31 | 0.31 | 0.00 |
| 11 | Reading | Read an Informational Passage | 0.28 | 0.37 | 0.09 | 0.34 | 0.14 | -0.20 |
| 11 | Reading | Read an Informational Passage | 0.61 | 0.66 | 0.05 | 0.53 | 0.50 | -0.02 |
| 11 | Reading | Read an Informational Passage | 0.51 | 0.57 | 0.06 | 0.52 | 0.40 | -0.13 |
| 11 | Reading | Read an Informational Passage | 0.34 | 0.36 | 0.02 | 0.28 | 0.26 | -0.03 |
| 11 | Reading | Read an Informational Passage | 0.64 | 0.55 | -0.09 | 0.50 | 0.64 | 0.14 |
| 11 | Reading | Read an Informational Passage | 0.58 | 0.48 | -0.10 | 0.54 | 0.62 | 0.08 |
| 11 | Writing | Describe a Picture | 0.76 | 0.79 | 0.03 | 0.56 | 0.58 | 0.02 |
| 11 | Writing | Describe a Picture | 0.84 | 0.84 | 0.00 | 0.51 | 0.53 | 0.02 |
| 11 | Writing | Write About an Experience | 0.75 | 0.73 | -0.02 | 0.61 | 0.82 | 0.20 |
| 11 | Writing | Write About Academic Information | 0.75 | 0.69 | -0.06 | 0.30 | 0.56 | 0.26 |
| 11 | Writing | Write About Academic Information | 0.64 | 0.67 | 0.03 | 0.70 | 0.81 | 0.10 |
| 11 | Writing | Justify an Opinion | 0.61 | 0.63 | 0.02 | 0.77 | 0.83 | 0.06 |

Table A.3 Written Language Kindergarten In Person Raw Scores

|  |  |  |  |
| --- | --- | --- | --- |
| Data Set | Reading Raw Score | Writing Raw Score | Number of Students |
| 1 | 0 | 4 | 1 |
| 2 | 2 | 1 | 1 |
| 3 | 3 | 4 | 1 |
| 4 | 3 | 12 | 1 |
| 5 | 4 | 0 | 1 |
| 6 | 4 | 2 | 1 |
| 7 | 5 | 3 | 1 |
| 8 | 5 | 5 | 1 |
| 9 | 5 | 11 | 1 |
| 10 | 6 | 2 | 1 |
| 11 | 6 | 6 | 1 |
| 12 | 6 | 7 | 1 |
| 13 | 6 | 8 | 2 |
| 14 | 7 | 0 | 1 |
| 15 | 7 | 1 | 3 |
| 16 | 7 | 3 | 2 |
| 17 | 7 | 5 | 2 |
| 18 | 7 | 6 | 1 |
| 19 | 7 | 9 | 1 |
| 20 | 8 | 0 | 1 |
| 21 | 8 | 1 | 1 |
| 22 | 8 | 3 | 1 |
| 23 | 8 | 5 | 3 |
| 24 | 8 | 9 | 3 |
| 25 | 8 | 11 | 1 |
| 26 | 8 | 12 | 1 |
| 27 | 9 | 0 | 1 |
| 28 | 9 | 1 | 1 |
| 29 | 9 | 2 | 2 |
| 30 | 9 | 3 | 2 |
| 31 | 9 | 4 | 1 |
| 32 | 9 | 5 | 1 |
| 33 | 9 | 6 | 4 |
| 34 | 9 | 7 | 1 |
| 35 | 9 | 8 | 5 |
| 36 | 9 | 10 | 1 |
| 37 | 9 | 11 | 3 |
| 38 | 9 | 12 | 1 |

Table A.3 *(continuation one)*

|  |  |  |  |
| --- | --- | --- | --- |
| Data Set | Reading Raw Score | Writing Raw Score | Number of Students |
| 39 | 10 | 2 | 1 |
| 40 | 10 | 3 | 2 |
| 41 | 10 | 4 | 1 |
| 42 | 10 | 5 | 2 |
| 43 | 10 | 6 | 1 |
| 44 | 10 | 7 | 1 |
| 45 | 10 | 8 | 4 |
| 46 | 10 | 10 | 2 |
| 47 | 10 | 11 | 2 |
| 48 | 10 | 12 | 3 |
| 49 | 11 | 1 | 1 |
| 50 | 11 | 2 | 1 |
| 51 | 11 | 3 | 2 |
| 52 | 11 | 6 | 2 |
| 53 | 11 | 7 | 4 |
| 54 | 11 | 8 | 1 |
| 55 | 11 | 9 | 2 |
| 56 | 11 | 10 | 3 |
| 57 | 11 | 11 | 1 |
| 58 | 11 | 12 | 4 |
| 59 | 12 | 5 | 1 |
| 60 | 12 | 7 | 2 |
| 61 | 12 | 8 | 1 |
| 62 | 12 | 9 | 3 |
| 63 | 12 | 10 | 1 |
| 64 | 12 | 11 | 2 |
| 65 | 12 | 12 | 12 |
| 66 | 13 | 2 | 2 |
| 67 | 13 | 3 | 2 |
| 68 | 13 | 4 | 1 |
| 69 | 13 | 5 | 1 |
| 70 | 13 | 6 | 1 |
| 71 | 13 | 8 | 1 |
| 72 | 13 | 9 | 1 |
| 73 | 13 | 10 | 2 |
| 74 | 13 | 11 | 10 |
| 75 | 13 | 12 | 11 |
| 76 | 14 | 2 | 1 |

Table A.3 *(continuation two)*

|  |  |  |  |
| --- | --- | --- | --- |
| Data Set | Reading Raw Score | Writing Raw Score | Number of Students |
| 77 | 14 | 3 | 1 |
| 78 | 14 | 5 | 1 |
| 79 | 14 | 7 | 5 |
| 80 | 14 | 8 | 1 |
| 81 | 14 | 9 | 1 |
| 82 | 14 | 10 | 1 |
| 83 | 14 | 11 | 8 |
| 84 | 14 | 12 | 18 |
| 85 | 15 | 4 | 1 |
| 86 | 15 | 5 | 1 |
| 87 | 15 | 6 | 1 |
| 88 | 15 | 7 | 2 |
| 89 | 15 | 8 | 2 |
| 90 | 15 | 9 | 3 |
| 91 | 15 | 10 | 3 |
| 92 | 15 | 11 | 9 |
| 93 | 15 | 12 | 38 |
| 94 | 16 | 4 | 1 |
| 95 | 16 | 9 | 2 |
| 96 | 16 | 10 | 3 |
| 97 | 16 | 11 | 10 |
| 98 | 16 | 12 | 76 |
| 99 | 17 | 3 | 1 |
| 100 | 17 | 9 | 1 |
| 101 | 17 | 10 | 1 |
| 102 | 17 | 11 | 10 |
| 103 | 17 | 12 | 139 |

Table A.4 Written Language Kindergarten Remote Raw Scores

|  |  |  |  |
| --- | --- | --- | --- |
| Data Set | Reading Raw Score | Writing Raw Score | Number of Students |
| 1 | 1 | 12 | 1 |
| 2 | 6 | 12 | 1 |
| 3 | 8 | 11 | 1 |
| 4 | 8 | 12 | 1 |
| 5 | 9 | 10 | 1 |
| 6 | 9 | 12 | 2 |
| 7 | 10 | 6 | 1 |
| 8 | 10 | 12 | 2 |
| 9 | 11 | 6 | 1 |
| 10 | 11 | 9 | 1 |
| 11 | 11 | 10 | 1 |
| 12 | 11 | 11 | 2 |
| 13 | 11 | 12 | 3 |
| 14 | 12 | 3 | 1 |
| 15 | 12 | 11 | 2 |
| 16 | 12 | 12 | 4 |
| 17 | 13 | 9 | 1 |
| 18 | 13 | 11 | 1 |
| 19 | 13 | 12 | 5 |
| 20 | 14 | 7 | 1 |
| 21 | 14 | 10 | 1 |
| 22 | 14 | 11 | 4 |
| 23 | 14 | 12 | 5 |
| 24 | 15 | 7 | 1 |
| 25 | 15 | 10 | 1 |
| 26 | 15 | 11 | 3 |
| 27 | 15 | 12 | 10 |
| 28 | 16 | 1 | 1 |
| 29 | 16 | 4 | 1 |
| 30 | 16 | 6 | 1 |
| 31 | 16 | 11 | 3 |
| 32 | 16 | 12 | 16 |
| 33 | 17 | 9 | 1 |
| 34 | 17 | 11 | 5 |
| 35 | 17 | 12 | 45 |