APPENDIX A

DETAILS OF STUDY DESIGN AND IMPLEMENTATION

A. STUDY DESIGN

This study was designed to estimate the impacts of four remedial reading programs using a scientifically rigorous experimental design involving both school-level and student-level randomization. The student-level randomization ensures that the estimates of the impacts of each of the interventions are obtained by comparing similar groups of students (the treatment and control groups), while the school-level randomization ensures that each of the interventions serves a similar mix of students, enabling comparison of the impacts of the four interventions.

We used this two-level random assignment design instead of a design where only schools were randomly assigned to one of the four interventions or a control condition for three primary reasons. First, randomizing students within schools gives us more power to detect significant impacts. Many more schools would have been required to obtain the same power if only schools were randomized. See Chapter II for more discussion of the statistical power of the study. Second, randomizing students within schools meant that we generated a control group from within the schools and so could provide the opportunity to participate in one of the interventions to at least some students in every school in the study. Preliminary discussions with AIU staff suggested that it would be much easier to gain the cooperation of the school districts and schools if we used our approach instead of the more straightforward one-level design, where none of the struggling readers in schools randomized to the control condition would receive any of the interventions. Third, we expect the distributions of student-level background covariates to be more similar between the treatment and control groups when students are randomized within schools than if only schools were randomized. This is due to the larger number of students relative to the number of schools. When randomizing a small number of units (e.g., schools), there is a higher probability that the treatment and control schools will be different on some background covariates, just by chance, than if there were a larger number of units randomized. In fact, in this study, by chance many of the smallest schools were randomized to the Wilson Reading condition. However, as discussed in Chapter II, this does not cause problems for the analysis because students were also randomized within schools.

The initial power analyses suggested that we could detect substantively meaningful impacts as statistically significant with a high probability when 40 schools were randomly assigned to one of four remedial reading interventions (Mathematica Policy Research, 2002). Each intervention would be implemented in 10 schools, and all treatment students within each school would receive the same intervention. Within each school and grade (grades 3 and 5), eligible students would then be randomly assigned to the treatment group to receive the intervention, or to the control group. The interventions would be delivered in instructional groups of three students, and the goal was to have two third-grade and two fifth-grade instructional groups within each school, with all four groups taught by one teacher. The expectation was that there would be approximately 10 eligible students in each school and grade, with six assigned to the treatment group and four to the control group. Once students were randomly assigned to the treatment and control groups within each school, we expected the program developers would form the instructional groups within the treatment group.

B. IMPLEMENTATION OF THE DESIGN

Figures A.1 through A.5 detail the implementation of the study design. Figure A.1 describes the school-level randomization, including the number of schools enrolled and randomized to each of the four interventions. As explained in Chapter II, it was necessary in some cases to form school units such that each school unit had instructional groups in both third and fifth grades. This first phase of randomization was done at the school-unit level.

Within the schools assigned to each of the interventions, we then used student-level randomization to assign students to either the treatment or control group. Figures A.2 through A.5 offer details on this randomization and the student-level data collection for the study.

Torgesen et al. (2006) analyzed principal and teacher survey data, data from the Common Core of Data (CCD), and Pennsylvania state assessment data to compare the schools in the study with other schools in the AIU and schools throughout Pennsylvania and the nation. They found that the study schools had a higher proportion of black students, a lower proportion of Hispanic students, and a lower proportion of students eligible for free or reduced-price lunches than the nation's schools. Compared with other schools in the AIU, the study schools had a higher proportion of black students, but similar proportions of Hispanic students and students eligible for free or reduced-price lunches. Torgesen et al. (2006) also found that the proportion of students scoring at or above the proficient level on the 2003 state fifth-grade reading assessment in the study schools was not significantly different from the proportion for all Pennsylvania schools. Furthermore, they found that across the groups of study schools assigned to the four interventions, the proportions of fifth-graders scoring proficient or higher in 2003 were not significantly different, suggesting that differences in instructional effectiveness that might occur across the interventions examined in this study should not be attributed to differences in effectiveness of the general reading instruction available at these schools.

Figure A-1

Consort Diagram for Schools in Study

All schools in the AIU encouraged to participate

Schools enrolled in study
(N=52)

School-units formed to have 3rd and 5th grades in each school-unit
(N=32 school-units)

School-units randomized within strata defined by percent eligible for free or reduced price lunch
(N=32)

Allocated to Failure-Free Reading
(N=8)
Stratum 1: N=1
Stratum 2: N=2
Stratum 3: N=2
Stratum 4: N=3

Allocated to Spell Read P.A.T.
(N=8)
Stratum 1: N=2
Stratum 2: N=2
Stratum 3: N=1
Stratum 4: N=3

Allocated to Wilson Reading
(N=8)
Stratum 1: N=2
Stratum 2: N=2
Stratum 3: N=2
Stratum 4: N=2

Allocated to Corrective Reading
(N=8)
Stratum 1: N=2
Stratum 2: N=1
Stratum 3: N=2
Stratum 4: N=3

1 school-unit (in Stratum 1) dropped out before knowing assignment

Schools identified potentially eligible students

Schools identified potentially eligible students

Schools identified potentially eligible students

Schools identified potentially eligible students

Figure A-2

Figure A-3

Figure A-4

Figure A-5

Figure A-2

Progression of Students in Schools Assigned to Failure Free Reading

School staff identified potential participants
(N = 438)

Excluded (N = 23)
    Parent refusal (N = 13)
    Transferred (N = 5)
    Other (N = 5)

Students Screened (Passive Consent)
(N = 415)

Excluded (N = 185)
    Ineligble (N = 100)
    Did not consent (N = 83)
    Other (N = 2)

Baseline data collected
    Student tests (N = 226)
    Parent survey (N = 223)
    Teacher survey (N = 225)

Not randomized (N=4)

Randomized
(N = 226)

Allocated to Treatment Group (N = 118)
    Received intervention (N = 116)
        Dropped out after 1st week (N = 0)
    Did not receive intervention (dropped out in 1st week) (N = 2)

Allocated to Control Group (N = 106)
    Received Intervention (N = 0)
    Did not receive intervention (N = 106)

Withdrew from the study (N = 1)
    First week (N = 0)
    After first week (N = 1)

Withdrew from the study (N = 1)
    First week (N = 0)
    After first week (N = 1)

PSSA data collection
    Student tests (N = 115)

2nd follow-up data collection
    Student tests (N = 113)
    Teacher survey (N = 115)
    School records (N = 108)

PSSA data collection
    Student tests (N = 102)

2nd follow-up data collection
    Student tests (N = 104)
    Teacher survey (N = 104)
    School records (N = 96)

In analysis sample
(N = 115)

In analysis sample
(N = 113)

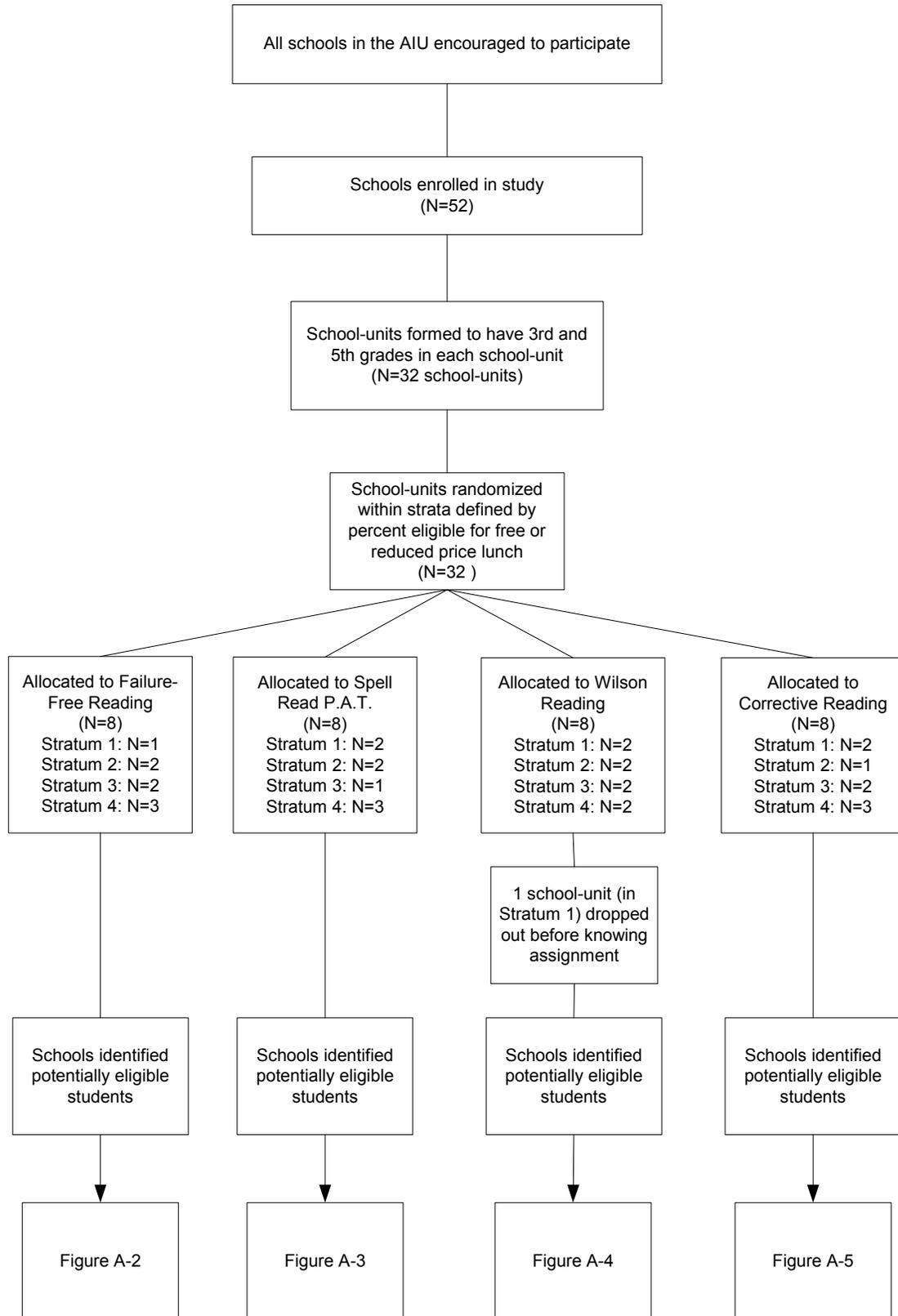In analysis sample
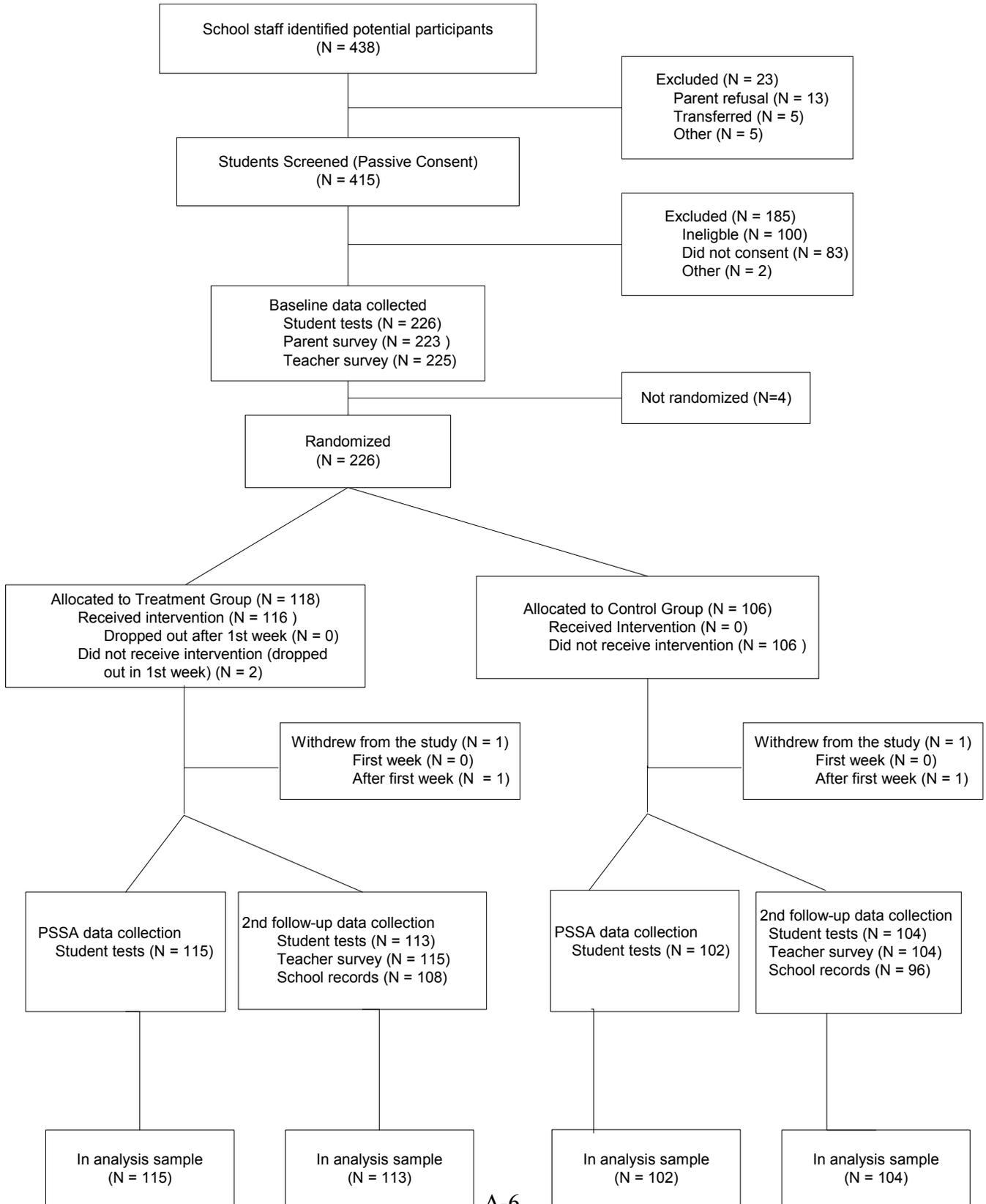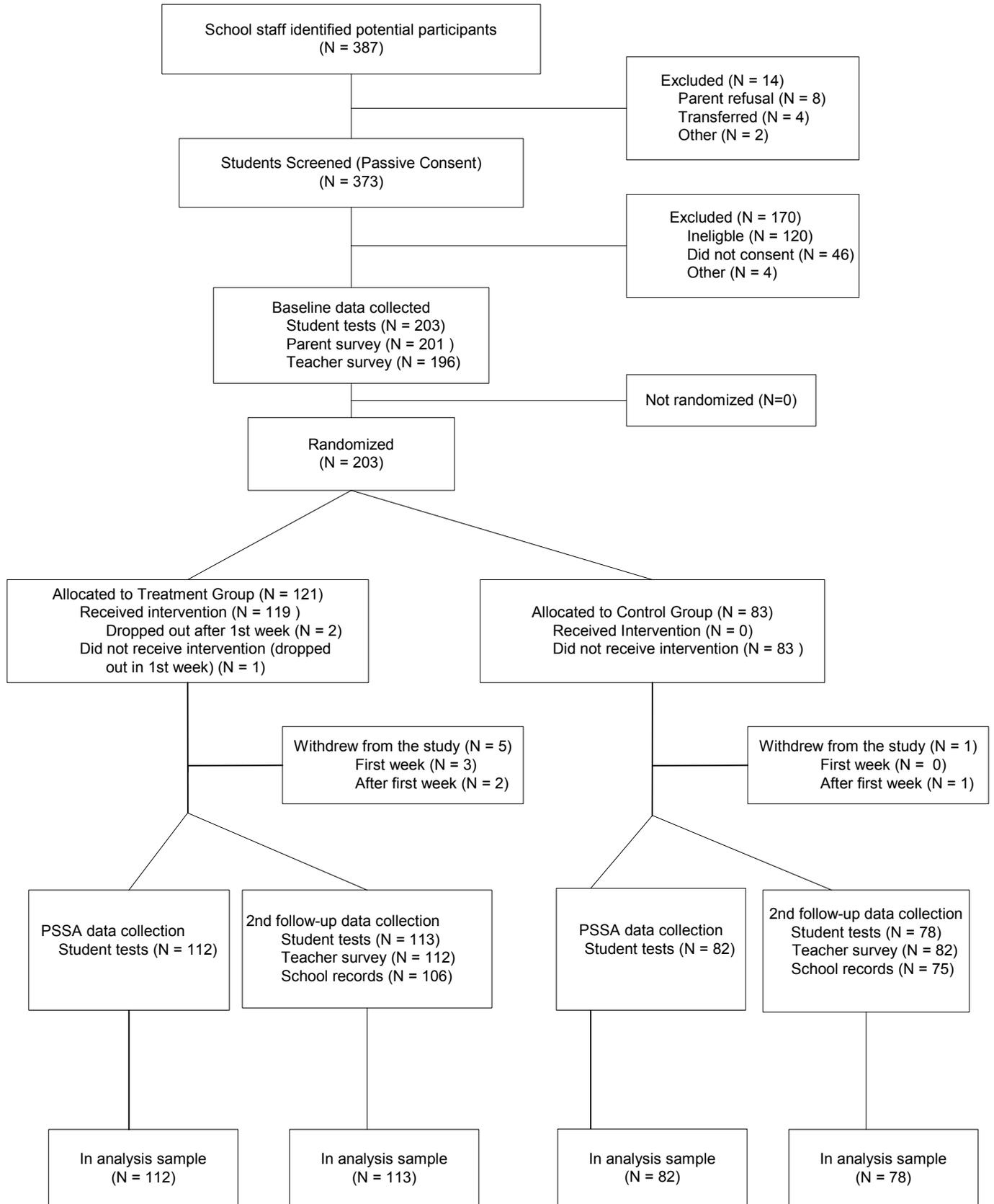(N = 102)

In analysis sample
(N = 104)

A-6

Figure A-3

Progression of Students in Schools Assigned to Spell Read P.A.T.

Figure A-4

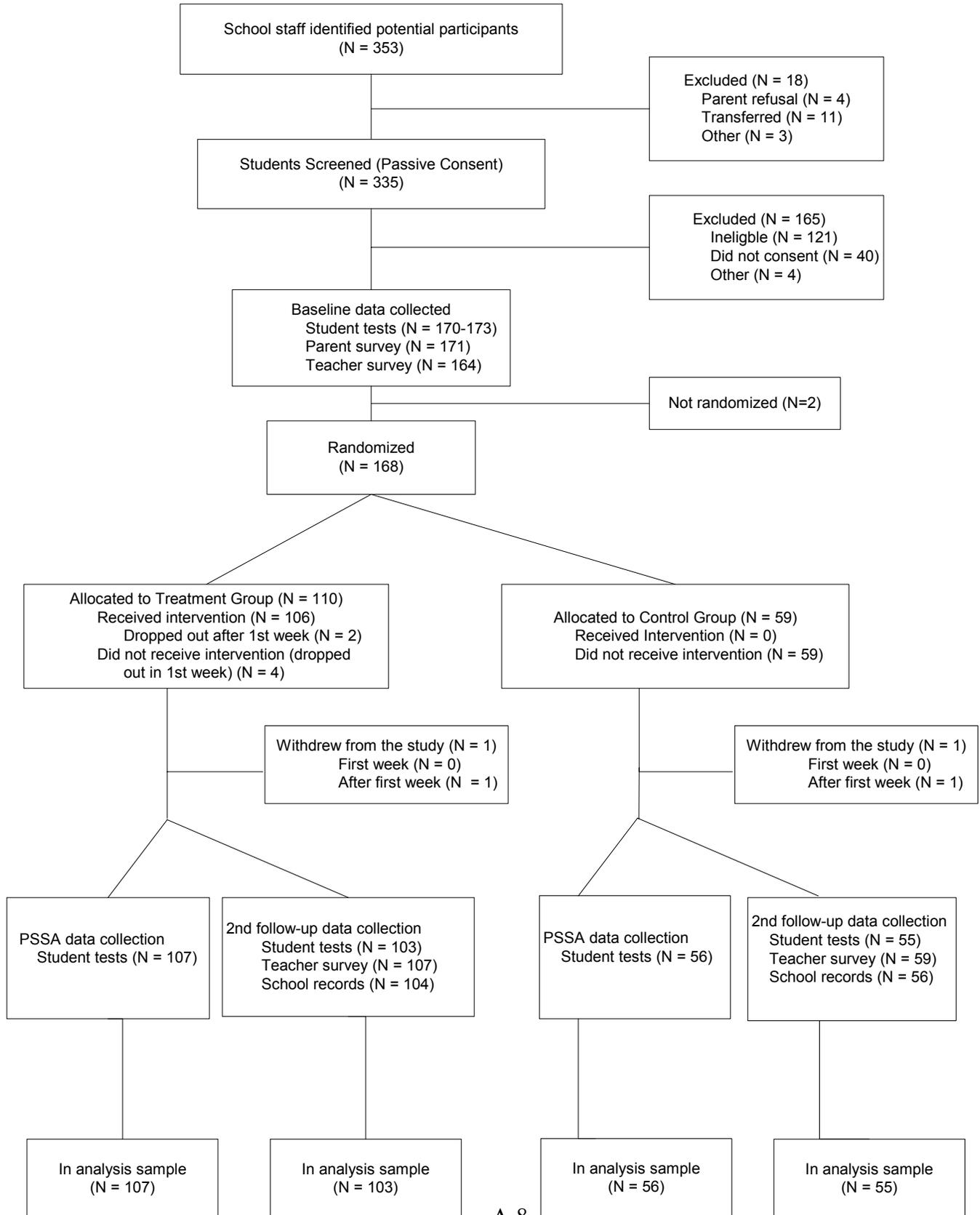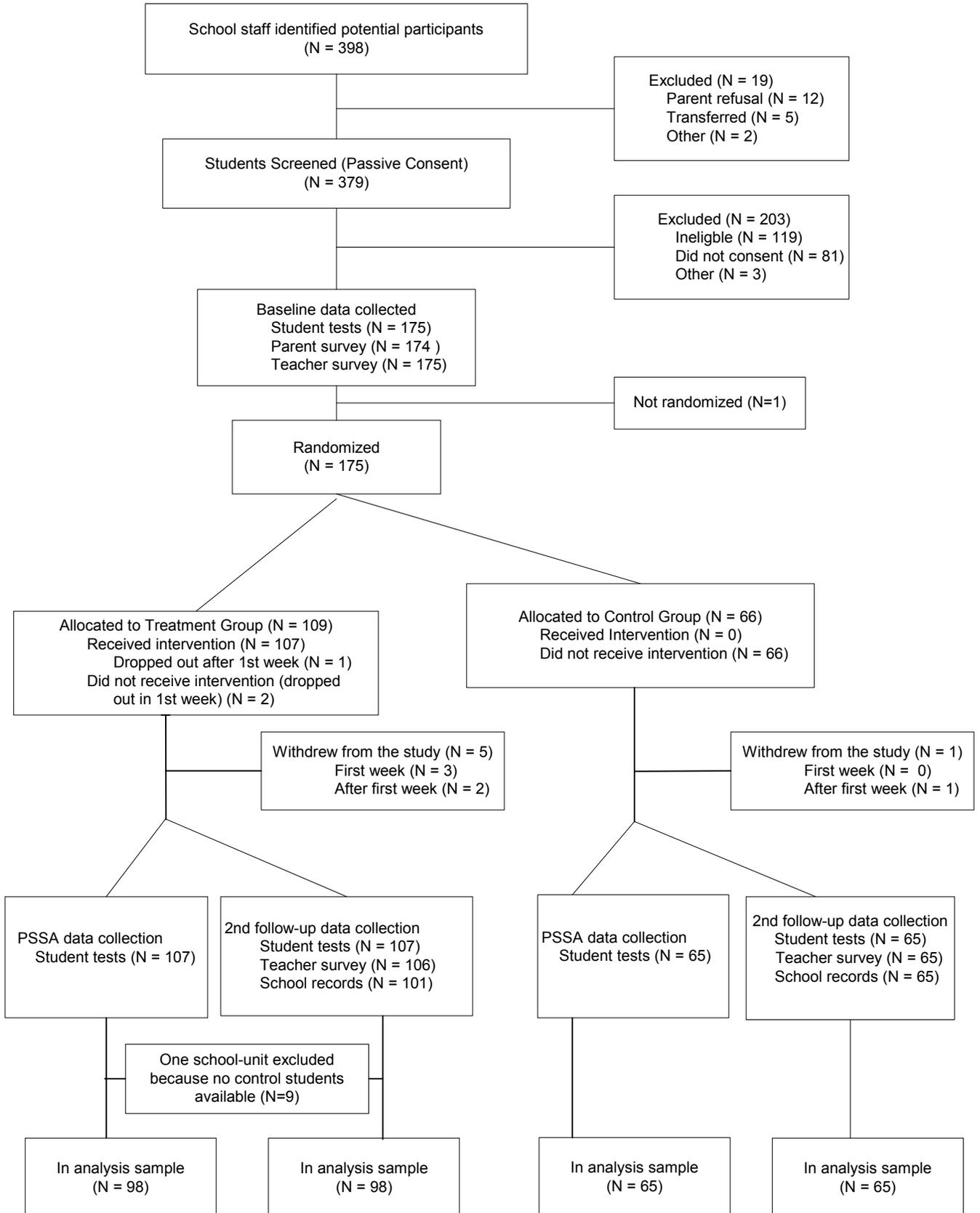Progression of Students in Schools Assigned to Wilson Reading

School staff identified potential participants
(N = 353)

Excluded (N = 18)
    Parent refusal (N = 4)
    Transferred (N = 11)
    Other (N = 3)

Students Screened (Passive Consent)
(N = 335)

Excluded (N = 165)
    Ineligble (N = 121)
    Did not consent (N = 40)
    Other (N = 4)

Baseline data collected
    Student tests (N = 170-173)
    Parent survey (N = 171)
    Teacher survey (N = 164)

Not randomized (N=2)

Randomized
(N = 168)

Allocated to Treatment Group (N = 110)
    Received intervention (N = 106)
    Dropped out after 1st week (N = 2)
    Did not receive intervention (dropped
    out in 1st week) (N = 4)

Allocated to Control Group (N = 59)
    Received Intervention (N = 0)
    Did not receive intervention (N = 59)

Withdrew from the study (N = 1)
    First week (N = 0)
    After first week (N = 1)

Withdrew from the study (N = 1)
    First week (N = 0)
    After first week (N = 1)

PSSA data collection
    Student tests (N = 107)

2nd follow-up data collection
    Student tests (N = 103)
    Teacher survey (N = 107)
    School records (N = 104)

PSSA data collection
    Student tests (N = 56)

2nd follow-up data collection
    Student tests (N = 55)
    Teacher survey (N = 59)
    School records (N = 56)

In analysis sample
(N = 107)

In analysis sample
(N = 103)

In analysis sample
(N = 56)

In analysis sample
(N = 55)

A-8

Figure A-5

Progression of Students in Schools Assigned to Corrective Reading

School staff identified potential participants
(N = 398)

Excluded (N = 19)
    Parent refusal (N = 12)
    Transferred (N = 5)
    Other (N = 2)

Students Screened (Passive Consent)
(N = 379)

Excluded (N = 203)
    Ineligble (N = 119)
    Did not consent (N = 81)
    Other (N = 3)

Baseline data collected
    Student tests (N = 175)
    Parent survey (N = 174)
    Teacher survey (N = 175)

Not randomized (N=1)

Randomized
(N = 175)

Allocated to Treatment Group (N = 109)
    Received intervention (N = 107)
    Dropped out after 1st week (N = 1)
    Did not receive intervention (dropped
    out in 1st week) (N = 2)

Allocated to Control Group (N = 66)
    Received Intervention (N = 0)
    Did not receive intervention (N = 66)

Withdrew from the study (N = 5)
    First week (N = 3)
    After first week (N = 2)

Withdrew from the study (N = 1)
    First week (N = 0)
    After first week (N = 1)

PSSA data collection
    Student tests (N = 107)

2nd follow-up data collection
    Student tests (N = 107)
    Teacher survey (N = 106)
    School records (N = 101)

PSSA data collection
    Student tests (N = 65)

2nd follow-up data collection
    Student tests (N = 65)
    Teacher survey (N = 65)
    School records (N = 65)

One school-unit excluded
because no control students
available (N=9)

In analysis sample
(N = 98)

In analysis sample
(N = 98)

In analysis sample
(N = 65)

In analysis sample
(N = 65)

A-9

APPENDIX B

DATA COLLECTION

This appendix describes the second follow-up data collection activities, which occurred in the spring semester of the 2004-05 school year. These activities included administering student reading tests, conducting teacher surveys, and extracting school records data. The appendix also describes our approach to gathering Pennsylvania System of School Assessment (PSSA) scores from the 2003-04 school year. Mathematica Policy Research (MPR) staff were responsible for administering the standardized reading tests. Personnel from the Allegheny Intermediate Unit (AIU) worked with MPR staff to coordinate all data collection activities within the AIU, including collecting teacher surveys, school records data, and PSSA scores.[1]

A. ADMINISTERING STANDARDIZED TESTS

Retired and substitute teachers were hired and trained to administer the standardized tests to students. Schools and districts identified former teachers or current substitute teachers to serve as test administrators in their schools, and some of those identified proposed other teachers to cover sites from which we had received no recommendations. Training lasted two days and focused predominantly on test administration. Other topics included an overview of the study, contacting and interacting with schools, scheduling tests, and arranging logistics. In addition to practice time incorporated in the training, testers were asked to conduct a "dress rehearsal" with a child one evening and bring questions to the next day's training. Many staff from the first year returned to administer tests in the second year. They participated in a separate, half-day, in-person, refresher training.

Testers were provided a toll-free telephone number to call if they had any questions once they began administering tests. Each also audiotaped one child's test administration session. The reading specialist who conducted the test administration training sessions reviewed the tapes and provided feedback. General reminders were e-mailed to all testers, and specific comments were discussed by telephone with individual testers.

Testers administered a battery of standardized tests about one month before the end of the 2004-05 school year. (Table II.8 lists the specific tests administered.) Tests were administered to both intervention and control students, including students who had transferred within a 100-mile radius of their original school.

The total test battery administration time was about one hour—half an hour for the individually administered tests and half an hour for the one group-administered test—but test times varied by student. When scheduling sessions, test administrators allowed for time beyond the actual test time to set up and put away materials for each test; to take students from their classroom to the testing room and then to return them to their classroom; and to calculate raw test scores. Testers permitted students to take a break between tests as needed.[2]

---

[1] As described in Torgesen et al. (2006), we obtained parental consent and student assent at the beginning of the study.

[2] Testers were not blind to the treatment/control status of sample members. Furthermore, although some of the school staff who assisted in extracting school records data and obtaining PSSA scores (as explained below) were blind to treatment/control status, not all of them were. Ensuring that all data collectors were blind to this information would have made the collection of data substantially more disruptive and burdensome for schools and students. In addition to this consideration, there was no apparent source of bias or other incentive to motivate intentional misreporting or

B. TEACHER SURVEYS

Near the end of the 2004-05 school year, reading/language arts teachers completed a survey for each student in the study. The survey collected information on the student's reading instruction, including the number of students in the child's reading group(s), the types of teachers providing reading instruction, and the amount of instruction received; pull-out reading instruction (curriculum areas missed, ways in which the child made up missed content); tutoring outside of school hours; views on the student's reading achievement (overall reading skills, improvement in reading, reading problems); and the frequency with which the student displayed various behaviors related to maintaining attention, adaptability, and social skills. The survey is shown in Appendix M.

The teacher surveys were disseminated to a school liaison (often the intervention teacher from the 2003-04 school year), who distributed them to and collected them from the students' reading/language arts teachers. For students who had transferred to non-AIU schools, teacher surveys were mailed to the child's school with a postage-paid, self-addressed return envelope. Telephone prompts were made to those schools to encourage the teachers to complete the survey.

C. EXTRACTING RECORDS DATA

At the conclusion of the 2004-05 academic year, school liaisons abstracted data to complete each student records form. These forms collected information on the student's enrollment, attendance, and suspensions; characteristics such as limited English proficiency, eligibility for the free or reduced-price lunch program, and disabilities; reading services; Individual Education Plan (IEP) or Service Agreement specifications; grade promotion and retention; course grades; and reading and math standardized test scores (Appendix N). As with the teacher surveys, student records forms for transfer students were mailed to the child's school with a postage-paid, self-addressed return envelope. Telephone prompts were made to those schools to encourage completion of the student records form.

D. OBTAINING PSSA SCORES

The PSSA is a standards-based, criterion-referenced assessment used to measure the attainment in certain areas of each Pennsylvania student in particular grades (http://www.pde.state.pa.us/a_and_t/site/default.asp). In the 2003-04 school year, students in grades three and five completed the PSSA in reading and math. (Fifth-grade students were also assessed in writing.) School liaisons entered raw and scaled scores, standard errors, and percentile scores in either an electronic or a hard copy form (Appendix K).

(continued)

manipulation of the data, and each data collector was responsible for obtaining data for only a fairly small proportion of all sample members, minimizing the likely effect of any problem that arose. Reviews of completed data collection instruments as they were received and subsequent analyses of data did not reveal any evidence of misreporting or other systematic problem. With respect to the school records data, the only variable that was used in our impact analyses was the indicator of the student's eligibility for free or reduced-price lunch.

APPENDIX C

WEIGHTING ADJUSTMENTS AND MISSING DATA

A.   BASIC WEIGHTING ADJUSTMENTS

Weighting adjustments are utilized to ensure that both the treatment group and the control group reflect the population of students in the study. The weighting is done within strata (the groups within which the student-level random assignment was done); these are generally defined by school and grade level.

This appendix describes the construction of baseline weights and "base" follow-up weights, and describes how the latter are adjusted to account for missing data. The study weights students within each randomization stratum up to the number of treatment slots that were originally allocated to that stratum.   These weights approximately correspond to a situation where schools receive funding proportional to enrollment.

1.   Construction of Baseline and "Base" Follow-Up Weights

The first step in calculating the weights was to create the set of counts to which we want to weight the data. These counts are the control totals to which the weights within each randomization stratum sum: the number of treatment slots originally allocated to each randomization stratum.

The first set of weights constructed are termed "baseline weights"; these weight all students in the study at baseline (the beginning of the school year), and weight treatment and control students together so that all students within a randomization stratum receive the same weight.   The total weight within each stratum equals the control total for that stratum.   For example, if the number of treatment slots in a stratum was six and there were nine consenting students (thus nine treatment and control students, combined), each student in that stratum would have a baseline weight of 6/9, or 2/3.

The next set of weights constructed are termed "preliminary base" weights for impact analyses. These are similar to the baseline weights, except that treatment and control students within each randomization stratum are weighted separately to the control total for that stratum. With a few exceptions (described below), the weight for each treatment student is the total divided by the number of treatment students, and the weight for each control student is the total divided by the number of control students. Continuing the example from the previous paragraph, if there were six treatment students and three control students in the stratum, then each treatment student would receive a base follow-up weight of 6/6, or 1, and each control student would receive a base follow-up weight of 6/3, or 2. Thus, the sum of the weights in each group within the stratum equals six, the number of treatment slots. This ensures that the treatment and control groups are weighted up to the same totals.

An initial complication in calculating these preliminary base weights is that of siblings where both siblings were in the evaluation sample. Generally, all students within a randomization stratum had the same probability of receiving treatment. However, siblings were randomized together to ensure that both siblings received the same assignment, avoiding situations in which one sibling was assigned to be in the treatment group while the other was assigned to be in the control group. This is not a problem if the two siblings were in the same randomization stratum or if the two strata had the same treatment probability.   However, if two siblings were from two different randomization strata with different probabilities (because of different ratios of eligible students to treatment slots in the two strata), the treatment probability for the sibling pair was the average of that from the two strata, and thus the siblings had a different probability of treatment assignment than the other students in the two strata. Thus, for 10 of the 79 randomization strata, slight adjustments were made to the weights to account for these differences in the treatment probabilities. These adjustments ensured that the ratio of the weight for a sibling to the weight for a non-sibling equaled the ratio of the inverse treatment probabilities for siblings and non-siblings.

B. WEIGHTING ADJUSTMENTS FOR MISSING DATA

1. Overview of Missing Data Challenges

The preliminary base weights described above were modified slightly to account for missing data on the tests used to assess reading ability and estimate impacts. There are two types of missing data: unit nonresponse, and item nonresponse. In this report, we estimate impacts on test scores from seven reading tests that we administered for the evaluation and test scores from reading and mathematics tests administered by AIU schools as part of the Pennsylvania System of School Assessment (PSSA). We were not able to administer the seven reading tests to all students in our evaluation sample, and we were not able to collect PSSA test scores for all of the students. The patterns of missing data for our tests and the PSSA tests, however, are different. There are 34 unit nonrespondents for our tests and 26 unit nonrespondents for the PSSA tests—in both instances, less than five percent of the evaluation sample. These students are unit nonrespondents because we are missing scores for all seven of our tests and both PSSA tests, respectively. We create separate analysis samples—one with respondents for our tests and one with respondents for the PSSA tests—and adjust for nonresponse separately (though using the same methodology).

In contrast to unit nonresponse, which occurs, for example, when there are missing scores for all seven of our reading tests, item nonresponse occurs when we are missing at least one score, but not all scores. On our baseline tests, there are three students missing one test score (each for a different test). For our second follow-up tests, which were administered one year after the interventions ended, two students are missing scores on one test, and another student is missing a score on a different test. Because there is so little item nonresponse—we are missing less than one-half of one percent of test scores—item nonrespondents are dropped from item-specific analyses without further adjustment. This should have a negligible effect on estimates.

As noted above, we use weighting adjustments to compensate for unit nonresponse. There are three types of unit nonresponse for which we utilize weighting adjustments: school by grade randomization strata in which there were no control students observed, a school unit for which there were no control students observed at all, and student unit nonresponse. The weighting adjustment method used weights up students with observed test scores who look similar to the students whose test scores are missing. Specifically, for unit nonresponse, we reassigned nonrespondent weights to other students in the same grade, blocking stratum (the strata within which the school-level randomization was done, defined by the percentage of students eligible for free or reduced-price lunch), and intervention using matching techniques that weight up other students who look similar to the student with missing information. These three types of missing data and the weighting adjustments are described further below.

2. Strata With No Control Students

Because the study was designed such that no intervention slots would go unfilled if there were enough consenting students to fill out an instructional group, in some randomization strata there were insufficient numbers of consenting students to assign any students to the control group. For example, if there were just three consenting students in a grade level within a school, all three would be assigned to receive the treatment and there would be no students assigned to the control group in that school and grade. There are seven such randomization strata.

For each of these strata, we weight up control students in other strata who look like the missing students. This similarity is defined as being in the same blocking stratum and grade, same race, gender, and income (when possible), and similar baseline test scores.[3,4] Because there is no information available on the missing control students, and because at baseline the treatment and control students within a randomization stratum should be only randomly different from one another, we do this by finding control students who look similar to the treatment students in the strata with missing control students.

Priority is given to close matching on the baseline test scores because they should be highly predictive of the outcome test scores. In addition, the matching is generally done so that exact matches on gender, race, and family income categories are obtained when possible—for example, an African American male student in a family with income over $30,000 is matched to other African American male students with family incomes over $30,000 and similar baseline test scores. In this way, control students with the same race, gender, and income level, and similar baseline test scores are weighted up to account for the missing control students in the strata with no control students. The weighting is done inversely proportional to the Mahalanobis distance, so that students with test scores more similar to those of the target are given greater weight than students with test scores less similar.

For example, suppose a stratum with no control students had three treated students, and that the control total for the weighting is six (so that the sum of the weights in the treatment and control groups in that stratum should each equal six). Each treatment student will get a weight of two. In addition, matches will be found for each treatment student in the stratum. Suppose treatment students one and two each have one match, while treatment student three has two matches, with Mahalanobis distances (D) of two and four, respectively. Each of the matches for treatment students one and two will have their weights increased by $w_k = 2 * \frac{1/D}{1/D} = 2$. For treatment student three, the two matches will have their weights increased by $w_1 = 2 * \frac{1/2}{1/2 + 1/4} = \frac{4}{3}$ and $w_2 = 2 * \frac{1/4}{1/2 + 1/4} = \frac{2}{3}$, respectively. Note that the sum of these weight increases is 2+2+4/3+2/3=6, which is the correct control total for that stratum.

3. School Unit With No Control Students

There is one school unit for which there are no control students observed at all, in either third or fifth grade. This case is more complicated than the randomization strata with no control students; it is not possible to estimate school unit parameters in the analysis model when there are no control students observed in an entire school unit. We thus drop this school-unit from the analysis, but weight up other students who look similar to the students in that school unit, to account for their being dropped. This

---

[3] We define "similar" test scores as having a small Mahalanobis distance, where the Mahalanobis distance between the test scores of student i and student j is defined as $(x_i - x_j)'\Sigma^{-1}(x_i - x_j)$, where x consists of baseline values of the nine test scores of primary interest (Aimsweb standard score, GRADE standardized score on passage comprehension, TOWRE 45-second phonemic decoding and awareness standard score, TOWRE 45-second sight word efficiency standard score, WRM-R word identification grade-based standard score, WRM-R word attack grade-based standard score, WRM-R passage comprehension grade-based standard score, WJIII calculation standard score, and WJIII spelling standard score), and Σ is the covariance matrix of these nine test scores in the full sample.

[4] There is no baseline measure for the PSSA available. For the purpose of computing the Mahalanobis distances, we used the baseline scores from the other tests.

will preserve the balance achieved through randomization; simply dropping the students from that school unit may degrade such balance.

A weighting adjustment similar to that for the strata with no control students was used here, except that in addition to control matches for the missing control students in this school unit, as described above, we also found treatment matches for the dropped treatment students. The method is similar to the procedure described above, except that the potential matches are restricted to treatment students in the same blocking stratum, grade, and intervention.

4. Nonresponse at Follow-Up

A final source of missing data is student-level (unit) nonresponse. An approach similar to that described above is used to adjust for this nonresponse, where students in the same treatment group, blocking strata, and grade, with the same age, race, family income, and similar baseline test scores, are weighted up to account for the unit nonrespondents. As noted above, we perform separate weight adjustments for our two analysis samples—one sample with the students who took our seven reading tests and one with the students for whom we obtained PSSA test scores.

C. IMPUTATION OF FREE OR REDUCED PRICE LUNCH STATUS

For most variables used in the analyses there was very little item-nonresponse. However, 24 students (3 percent) had a missing value for the variable indicating free or reduced-price lunch status, as reported in the school records. Because this variable was used to define a key subgroup of interest and because we had supplementary information from the parent survey, we used hot-deck imputation (Little and Rubin, 2002) to impute free or reduced-price lunch eligibility for those 24 students. Specifically, for students with missing free or reduced-price lunch eligibility, we imputed their eligibility by randomly selecting one student with the same family income level and family size (as reported in the parent survey) and imputing their observed eligibility status for the student with missing eligibility status. This enables the use of all students in the free or reduced-price lunch eligibility subgroup analyses.

APPENDIX D

DETAILS OF STATISTICAL METHODS

A. ESTIMATION OF IMPACTS

The impacts are defined as the regression-adjusted difference in the average achievement scores for the treatment and the control groups. From the hierarchical linear model (HLM), we estimate impacts for each of the four interventions.[5] We also estimate the impact of being assigned to any of the interventions, denoted as the pooled intervention impact (ABCD), as the average of these four intervention impacts.[6]

As discussed in Chapter IV, we can obtain our impact estimates from our model by substituting in the level-two equations (Equations (IV.2) in Chapter IV) into the level-one equation (Equation (IV.1) in Chapter IV) and considering this combined equation for different profiles of students and school units.[7] We assume that each blocking stratum contributes equally to the impacts and construct impacts using an equally weighted average blocking effect. [8]

In addition to estimating impacts for the all students in the third-grade cohort and all students in the fifth-grade cohort, we estimate impacts for subgroups of students within each grade cohort. Being able to estimate impacts for subgroups and to test for differences in impacts among them allows for potentially better targeting of the interventions to, for example, students with especially low phonemic decoding skills. To estimate subgroup impacts, we modify the model specification found in Chapter IV to allow for a general subgroup, S_{ij}, as described below.

---

[5] We used HLM 6 © software published by Scientific Software International, Inc. to obtain the HLM estimates. Parameter estimates are obtained using restricted maximum likelihood (REML), as discussed in Raudenbush and Bryk (2002).

[6] Since we assume that each intervention contributes equally to the pooled intervention impact, we use an equally weighted average of the individual interventions.

[7] For example, the equation pertaining to control students from school units assigned to intervention D is

$$Y_{ij} = \gamma_{00} + \gamma_{20} y_{oij}^* + (1/4)(\xi_{01} + \xi_{02} + \xi_{03}) + (1/4)\left[(\xi_{21} + \xi_{22} + \xi_{23}) y_{oij}^*\right] + error$$

And, the equation pertaining to treatments in intervention A is

$$Y_{ij} = \gamma_{00} + \gamma_{01} + \gamma_{10} + \gamma_{11} + (\gamma_{20} + \gamma_{21}) y_{oij}^* + (1/4)(\xi_{01} + \xi_{02} + \xi_{03}) + (1/4)(\xi_{11} + \xi_{12} + \xi_{13})$$
$$+ (1/4)\left[(\xi_{21} + \xi_{22} + \xi_{23}) y_{oij}^*\right] + error.$$

Similar equations can be obtained for each intervention and treatment or control group.

[8] An estimate of impacts for different profiles of students could include those students in a particular intervention within a particular blocking strata. Since we assume blocking strata contribute equally to the impact estimates, we capture the blocking effect by using an average blocking effect, where each blocking strata contributes equally. This average is an equally weighted average of impacts for interventions and combinations of interventions from each of the four blocking strata. For example, impacts for interventions A in blocking strata 1, 2, 3 and 4 are averaged so that each blocking strata contributes equally to impact of intervention A. Thus, for the third-grade cohort, the term $(1/4)(\hat{\xi}_{11} + \hat{\xi}_{12} + \hat{\xi}_{13})$ in Equation (IV.5) captures the blocking effect in intervention A impacts.

B. IMPACTS FOR GENERAL SUBGROUPS MEASURED AT THE INDIVIDUAL LEVEL

The model specification found in Chapter IV is modified as follows to allow for different impacts for a general subgroup, S_{ij}.

Level One: Student i within school unit j

$$y_{1ij} = \beta_{0j} + \beta_{1j}T_{ij} + \beta_{2j}y^*_{oij} + \beta_{3j}S_{ij} + \beta_{4j}T_{ij}S_{ij} + r_{ij}$$

Level Two: School Unit j

$$\beta_{kj} = \gamma_{k0} + \gamma_{k1}A_j + \gamma_{k2}B_j + \gamma_{k3}C_j + \sum_{l=1}^{3}\xi_{kl}P_{lj} + \mu_{kj},$$

for level-one parameter β_{kj}, $k = 0,\text{K},4$, where T_{ij} is as defined as in Chapter IV, and

$$S_{ij} = 1 \text{ if student } i \text{ in school-unit } j \text{ is in the subgroup, and}$$

$$S_{ij} = 0 \text{ if student } i \text{ in school-unit } j \text{ is not in the subgroup.}$$

C. ADJUSTMENT FOR MULTIPLE COMPARISONS

When estimating impacts for multiple outcomes and testing multiple interventions, there is a concern that some estimated impacts will be found to be significantly different from zero, even if there is actually no impact of the interventions (a "Type 1" error). In fact, even if there were no differences between the treatment and control groups, 5 percent of test statistics comparing the outcomes of the two groups would be expected to be significant at the 5 percent level. To address these types of concerns, methods have been developed that adjust the significance levels of tests to account for the number of tests being performed.

We implemented two of these methods that adjust significance levels to account for the multiple comparisons being performed. The first is the Bonferroni correction, which is a commonly used method that controls the familywise error rate, ensuring that the probability of making any Type 1 error—rejecting a null hypothesis that is in fact true—is at the designated level (e.g., 0.05). The drawback of the Bonferroni method is that it often has low statistical power. The second method we utilized was developed by Benjamini and Hochberg (BH) in 1995. The BH method controls the "false discovery rate," or the expected proportion of falsely rejected hypotheses (Type 1 errors). This procedure is more powerful than the Bonferroni method, and is particularly relevant for situations such as this, where interest is in assessing the impact of an intervention on multiple outcomes.

A key consideration in using these procedures is the grouping of tests; the adjustments are done within groups and they depend on the number of tests within each group. We use two groupings: the seven reading tests within each grade level and contrast examined (e.g., ABCD), and the two or three tests within each reading domain (phonemic decoding and awareness, reading accuracy and fluency, and reading comprehension), grade level, and contrast. We used these groupings because we are generally interested in statements such as "the four interventions combined increased phonemic decoding among students in the third-grade cohort," making statements separately for the third-and fifth-grade cohorts, and separately for each type of contrast (the four interventions, the three word-level interventions, and each of the four interventions individually).

Tables D.1 through D.20 show the significance levels with no adjustment (as presented in Chapter IV), adjusted using the Bonferroni method, and adjusted using the BH method, for the full sample and for each of the subgroups. The odd-numbered tables do the adjustments within tests grouped by type; the even-numbered tables do the adjustments within groups of all seven tests. The corresponding sample sizes are reported in Tables IV.1 to IV.10.

Table D.1

3rd and 5th Grade Cohorts: Comparison of Significance Levels, Tests Grouped by Type

Grade 3 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
TOWRE PDE	*	*	*	*	*	*				*	*	*	*	*	*	*		
Word Identification	*	*	*	*	*	*							*	*	*	*		
TOWRE SWE	*	*	*	*	*	*	*		*				*	*	*			
AIMSweb							*	*	*	*								
Passage Comprehension	*	*	*				*	*	*									
GRADE																		

Grade 5 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*							*	*	*			
TOWRE PDE				*	*	*												
Word Identification																		
TOWRE SWE										*	*	*						
AIMSweb	*			*														
Passage Comprehension																		
GRADE																		

* Impact is statistically significant at the 0.05 level.

Note: "Unadj." is unadjusted p-values, "Bonf." referes to p-values adjusted using Bonferroni procedure, "BH" refers to p-values adjusted using Benjamini and Hochberg (1995) procedure.

Table D.2

3rd and 5th Grade Cohorts: Comparison of Significance Levels, All Tests Grouped Together

Grade 3 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*	*	░	*	*	*	*	*	*	*	*	░	░
TOWRE PDE	*	*	*	*	*	*				*	*	*	*	*	*	*	░	░
Word Identification	*	*	*	*	*	*							*	*	*	*	░	░
TOWRE SWE	*	*	*	*	░	*	*	░	*				*	░	*			
AIMSweb							*	*	*	*	░	*						
Passage Comprehension	*	░	*				*	*	*									
GRADE																		

Grade 5 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*							*	*	*			
TOWRE PDE				*	░	░												
Word Identification																		
TOWRE SWE										*	*	*						
AIMSweb	*	░	░	*	░	░												
Passage Comprehension																		
GRADE																		

* Impact is statistically significant at the 0.05 level.
Note: "Unadj." is unadjusted p-values, "Bonf." referes to p-values adjusted using Bonferroni procedure, "BH" refers to p-values adjusted using Benjamini and Hochberg (1995) procedure.

Table D.3

3rd and 5th Grade Cohorts with Low Word Attack Scores: Comparison of Significance Levels, Tests Grouped by Type

Grade 3 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*				*	*	*						
TOWRE PDE	*	*	*	*						*	*	*						
Word Identification	*	*	*	*									*	*	*			
TOWRE SWE																		
AIMSweb							*	*	*	*								
Passage Comprehension																		
GRADE																		

Grade 5 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*				*	*	*	*	*	*	*	*	*
TOWRE PDE				*		*												
Word Identification																		
TOWRE SWE	*						*			*								
AIMSweb	*			*									*					
Passage Comprehension																		
GRADE	*			*	*	*							*	*	*			

* Impact is statistically significant at the 0.05 level.

Note: "Unadj." is unadjusted p-values, "Bonf." referes to p-values adjusted using Bonferroni procedure, "BH" refers to p-values adjusted using Benjamini and Hochberg (1995) procedure.

Table D.4

3rd and 5th Grade Cohorts with Low Word Attack Scores: Comparison of Significance Levels, All Tests Grouped Together

Grade 3 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	▓	▓				*	▓	▓						
TOWRE PDE	*	▓	*	*	▓	▓				*								
Word Identification	*	*	*	*	▓	▓							*	*	*			
TOWRE SWE																		
AIMSweb							*	▓	▓	*	▓	▓						
Passage Comprehension																		
GRADE																		

Grade 5 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*				*	*	*	*	▓	*	*	▓	▓
TOWRE PDE				*	▓	▓												
Word Identification																		
TOWRE SWE	*	▓	▓				*	▓	▓	*	▓	▓						
AIMSweb	*	▓	▓	*	▓	▓							*	▓	▓			
Passage Comprehension																		
GRADE	*	▓	▓	*	▓	*							*	*	*			

* Impact is statistically significant at the 0.05 level.
Note: "Unadj." is unadjusted p-values, "Bonf." referes to p-values adjusted using Bonferroni procedure, "BH" refers to p-values adjusted using Benjamini and Hochberg (1995) procedure.

D-9

Table D.5

3rd and 5th Grade Cohorts with High Word Attack Scores: Comparison of Significance Levels, Tests Grouped by Type

Grade 3 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*	*						*		*	*	*	*
TOWRE PDE	*	*	*	*	*	*				*	*	*	*	*	*	*	*	*
Word Identification	*	*	*	*	*	*	*		*				*	*	*			
TOWRE SWE	*		*	*		*	*	*	*				*	*	*	*	*	*
AIMSweb	*		*	*		*							*		*			
Passage Comprehension	*	*	*				*	*	*									
GRADE							*		*				*	*	*			

Grade 5 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack													*	*	*			
TOWRE PDE																		
Word Identification	*																	
TOWRE SWE										*	*	*						
AIMSweb																		
Passage Comprehension	*	*	*													*		*
GRADE							*	*	*							*		*

* Impact is statistically significant at the 0.05 level.
Note: "Unadj." is unadjusted p-values, "Bonf." referes to p-values adjusted using Bonferroni procedure, "BH" refers to p-values adjusted using Benjamini and Hochberg (1995) procedure.

Table D.6

3rd and 5th Grade Cohorts with High Word Attack Scores: Comparison of Significance Levels, All Tests Grouped Together

Grade 3 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*	*						*		*	*		*
TOWRE PDE	*	*	*	*	*	*				*	*	*	*	*	*	*		*
Word Identification	*	*	*	*	*	*	*		*				*		*			
TOWRE SWE	*	*	*	*	*	*	*	*	*				*		*	*	*	*
AIMSweb	*		*	*		*							*		*			
Passage Comprehension	*		*				*	*	*				*	*	*			
GRADE							*		*									

Grade 5 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack													*	*	*			
TOWRE PDE																		
Word Identification	*																	
TOWRE SWE										*	*	*						
AIMSweb																		
Passage Comprehension	*															*		
GRADE							*									*		

* Impact is statistically significant at the 0.05 level.

Note: "Unadj." is unadjusted p-values, "Bonf." referes to p-values adjusted using Bonferroni procedure, "BH" refers to p-values adjusted using Benjamini and Hochberg (1995) procedure.

Table D.7

3rd and 5th Grade Cohorts with Low Screening Peabody Picture Vocabulary Test Scores: Comparison of Significance Levels, Tests Grouped by Type

Grade 3 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*	*	*	*				*	*	*	*	[shaded]	*
TOWRE PDE	*	*	*	*	*	*							*	*	*	*	*	*
Word Identification													*	*	*			
TOWRE SWE																		
AIMSweb																		
Passage Comprehension																		
GRADE																		

Grade 5 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*							*	*	*			
TOWRE PDE	*	*	*	*	*	*												
Word Identification																		
TOWRE SWE	*	*	*	*	*	*				*	*	*				*	[shaded]	
AIMSweb																		
Passage Comprehension				*	[shaded]	*							*	*	*			
GRADE	*	*	*	*	*	*	*	[shaded]					*	*	*	*	[shaded]	

* Impact is statistically significant at the 0.05 level.
Note: "Unadj." is unadjusted p-values, "Bonf." referes to p-values adjusted using Bonferroni procedure, "BH" refers to p-values adjusted using Benjamini and Hochberg (1995) procedure.

Table D.8

3rd and 5th Grade Cohorts with Low Screening Peabody Picture Vocabulary Test Scores: Comparison of Significance Levels, All Tests Grouped Together

Grade 3 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*			*						*	*	*	*		
TOWRE PDE	*	*	*	*	*	*							*		*	*	*	*
Word Identification													*		*			
TOWRE SWE																		
AIMSweb																		
Passage Comprehension																		
GRADE																		

Grade 5 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*							*	*	*			
TOWRE PDE	*	*	*	*	*	*												
Word Identification																		
TOWRE SWE	*	*	*	*	*	*				*	*	*				*		
AIMSweb																		
Passage Comprehension				*									*	*	*			
GRADE	*		*	*			*						*		*	*		

* Impact is statistically significant at the 0.05 level.
Note: "Unadj." is unadjusted p-values, "Bonf." referes to p-values adjusted using Bonferroni procedure, "BH" refers to p-values adjusted using Benjamini and Hochberg (1995) procedure.

Table D.9

3rd and 5th Grade Cohorts with High Screening Peabody Picture Vocabulary Test Scores: Comparison of Significance Levels, Tests Grouped by Type

Grade 3 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*				*	*	*				*	*	*
TOWRE PDE	*	*	*	*	*	*				*	*	*						
Word Identification	*		*	*	*	*	*		*				*	*	*			
TOWRE SWE	*		*				*	*	*									
AIMSweb	*	*	*	*	*	*	*	*	*	*			*	*	*			
Passage Comprehension	*	*	*				*	*	*									
GRADE																		

Grade 5 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*		*	*	*	*				*	*	*						
TOWRE PDE	*	*	*	*	*	*				*	*	*						
Word Identification																		
TOWRE SWE										*								
AIMSweb	*	*	*	*	*	*										*	*	*
Passage Comprehension																		
GRADE																		

* Impact is statistically significant at the 0.05 level.

Note: "Unadj." is unadjusted p-values, "Bonf." referes to p-values adjusted using Bonferroni procedure, "BH" refers to p-values adjusted using Benjamini and Hochberg (1995) procedure.

Table D.10

3rd and 5th Grade Cohorts with High Screening Peabody Picture Vocabulary Test Scores: Comparison of Significance Levels, All Tests Grouped Together

Grade 3 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*				*	*	*				*		
TOWRE PDE	*	*	*	*	*	*				*	*	*						
Word Identification	*		*	*	*	*	*		*				*	*	*			
TOWRE SWE	*		*				*		*									
AIMSweb	*	*	*	*	*	*	*	*	*	*			*	*	*			
Passage Comprehension	*		*				*	*	*									
GRADE																		

Grade 5 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*			*		*				*	*	*						
TOWRE PDE	*			*	*	*				*	*	*						
Word Identification																		
TOWRE SWE										*								
AIMSweb	*	*	*	*	*	*										*	*	*
Passage Comprehension																		
GRADE																		

* Impact is statistically significant at the 0.05 level.
Note: "Unadj." is unadjusted p-values, "Bonf." referes to p-values adjusted using Bonferroni procedure, "BH" refers to p-values adjusted using Benjamini and Hochberg (1995) procedure.

Table D.11

3rd and 5th Grade Cohorts with Low Baseline Word Attack Scores and Low Screening Peabody Picture Vocabulary Test Scores: Comparison of Significance Levels, Tests Grouped by Type

Grade 3 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*										*		
TOWRE PDE																		
Word Identification																		
TOWRE SWE																		
AIMSweb	*			*	*	*							*	*	*			
Passage Comprehension																		
GRADE																		

Grade 5 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*				*	*	*				*	*	*
TOWRE PDE	*	*	*	*	*	*												
Word Identification																		
TOWRE SWE	*									*								
AIMSweb	*			*														
Passage Comprehension	*	*	*	*	*	*				*	*	*				*		
GRADE																		

* Impact is statistically significant at the 0.05 level.
Note: "Unadj." is unadjusted p-values, "Bonf." referes to p-values adjusted using Bonferroni procedure, "BH" refers to p-values adjusted using Benjamini and Hochberg (1995) procedure.

Table D.12

3rd and 5th Grade Cohorts with Low Baseline Word Attack Scores and Low Screening Peabody Picture Vocabulary Test Scores: Comparison of Significance Levels, All Tests Grouped Together

Grade 3 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack / TOWRE PDE	*	*	*	*	*	*										*		
Word Identification / TOWRE SWE / AIMSweb	*			*	*	*							*	*	*			
Passage Comprehension / GRADE																		

Grade 5 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*				*		*				*	*	*
TOWRE PDE	*		*	*		*												
Word Identification																		
TOWRE SWE	*									*		*						
AIMSweb	*			*														
Passage Comprehension / GRADE	*	*	*	*	*	*				*	*	*				*		

* Impact is statistically significant at the 0.05 level.

Note: "Unadj." is unadjusted p-values, "Bonf." referes to p-values adjusted using Bonferroni procedure, "BH" refers to p-values adjusted using Benjamini and Hochberg (1995) procedure.

D-17

Table D.13

3rd and 5th Grade Cohorts with Low Baseline Word Attack Scores and High Screening Peabody Picture Vocabulary Test Scores: Comparison of Significance Levels, Tests Grouped by Type

Grade 3 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack																*	*	*
TOWRE PDE										*	*	*						
Word Identification	*	▓	*										*	*	*			
TOWRE SWE																		
AIMSweb	*	*	*	*	▓	▓	*	▓	▓									
Passage Comprehension	*	▓	▓				*	*	*									
GRADE							*	*	*									

Grade 5 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*				*	*	*	*	*	*			
TOWRE PDE	*	*	*	*	*	*				*	*	*	*	▓	*	*	▓	▓
Word Identification							*	▓	▓									
TOWRE SWE	*	*	*	*	*	*				*	*	*						
AIMSweb													*	▓	▓			
Passage Comprehension																		
GRADE	*	*	*	*	*	*							*	*	*			

* Impact is statistically significant at the 0.05 level.
Note: "Unadj." is unadjusted p-values, "Bonf." referes to p-values adjusted using Bonferroni procedure, "BH" refers to p-values adjusted using Benjamini and Hochberg (1995) procedure.

Table D.14

3rd and 5th Grade Cohorts with Low Baseline Word Attack Scores and High Screening Peabody Picture Vocabulary Test Scores: Comparison of Significance Levels, All Tests Grouped Together

Grade 3 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack																*		
TOWRE PDE										*	*	*						
Word Identification	*												*	*	*			
TOWRE SWE																		
AIMSweb	*			*			*											
Passage Comprehension	*						*	*	*									
GRADE							*											

Grade 5 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*				*		*	*	*	*			
TOWRE PDE	*	*	*	*	*	*				*	*	*	*		*	*		
Word Identification							*											
TOWRE SWE	*		*	*		*				*	*	*						
AIMSweb													*		*			
Passage Comprehension																		
GRADE	*		*	*	*	*							*		*			

* Impact is statistically significant at the 0.05 level.

Note: "Unadj." is unadjusted p-values, "Bonf." referes to p-values adjusted using Bonferroni procedure, "BH" refers to p-values adjusted using Benjamini and Hochberg (1995) procedure.

Table D.15

3rd and 5th Grade Cohorts with High Baseline Word Attack Scores and High Screening Peabody Picture Vocabulary Test Scores: Comparison of Significance Levels, Tests Grouped by Type

Grade 3 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*	*			*	*	*	*		*	*	*	*
TOWRE PDE	*	*	*	*	*	*				*	*	*	*	*	*			
Word Identification	*			*	*	*										*	*	*
TOWRE SWE	*	*	*	*	*	*	*	*	*				*	*	*	*	*	*
AIMSweb													*	*	*			
Passage Comprehension	*	*	*				*	*	*									
GRADE																		

Grade 5 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack																		
TOWRE PDE	*			*														
Word Identification										*								
TOWRE SWE																		
AIMSweb																*	*	*
Passage Comprehension	*	*	*	*			*	*	*							*	*	*
GRADE																		

* Impact is statistically significant at the 0.05 level.
Note: "Unadj." is unadjusted p-values, "Bonf." referes to p-values adjusted using Bonferroni procedure, "BH" refers to p-values adjusted using Benjamini and Hochberg (1995) procedure.

D-20

Table D.16

3rd and 5th Grade Cohorts with High Baseline Word Attack Scores and High Screening Peabody Picture Vocabulary Test Scores: Comparison of Significance Levels, All Tests Grouped Together

Grade 3 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*	*			*	*	*	*			*	*	*
TOWRE PDE	*	*	*	*	*	*				*	*	*	*		*			
Word Identification	*			*		*										*	*	*
TOWRE SWE	*	*	*	*	*	*	*	*	*				*		*	*	*	*
AIMSweb													*	*	*			
Passage Comprehension	*	*	*				*	*	*									
GRADE																		

Grade 5 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*			*														
TOWRE PDE																		
Word Identification										*								
TOWRE SWE																		
AIMSweb																*	*	*
Passage Comprehension	*			*			*											
GRADE																*	*	*

* Impact is statistically significant at the 0.05 level.
Note: "Unadj." is unadjusted p-values, "Bonf." referes to p-values adjusted using Bonferroni procedure, "BH" refers to p-values adjusted using Benjamini and Hochberg (1995) procedure.

Table D.17

3rd and 5th Grade Cohorts Eligible for Free or Reduced-Price School Lunch: Comparison of Significance Levels, Tests Grouped by Type

Grade 3 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	▓	*	*	*	*				*	*	*			
TOWRE PDE	*	*	*	*	*	*				*	*	*	*	*	*			
Word Identification													*	*	*			
TOWRE SWE	*	▓	▓				*	▓	▓				*	▓	*			
AIMSweb																		
Passage Comprehension																		
GRADE																		

Grade 5 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*							*	*	*			
TOWRE PDE																*	*	*
Word Identification	*	*	*	*	*	*				*	*	*				*	*	*
TOWRE SWE																		
AIMSweb																		
Passage Comprehension	*	*	*	*	▓	▓	*	*	*	*	*	*						
GRADE																		

* Impact is statistically significant at the 0.05 level.
Note: "Unadj." is unadjusted p-values, "Bonf." referes to p-values adjusted using Bonferroni procedure, "BH" refers to p-values adjusted using Benjamini and Hochberg (1995) procedure.

Table D.18

3rd and 5th Grade Cohorts Eligible for Free or Reduced-Price School Lunch: Comparison of Significance Levels, All Tests Grouped Together

Grade 3 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*			*	*	*				*	*	*			
TOWRE PDE	*		*	*	*	*				*	*	*	*	*	*			
Word Identification													*	*	*			
TOWRE SWE	*						*						*					
AIMSweb																		
Passage Comprehension																		
GRADE																		

Grade 5 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*							*					
TOWRE PDE																*		*
Word Identification																		
TOWRE SWE	*	*	*	*	*	*				*	*	*				*	*	*
AIMSweb																		
Passage Comprehension																		
GRADE	*	*	*	*			*	*	*	*	*	*						

* Impact is statistically significant at the 0.05 level.
Note: "Unadj." is unadjusted p-values, "Bonf." referes to p-values adjusted using Bonferroni procedure, "BH" refers to p-values adjusted using Benjamini and Hochberg (1995) procedure.

Table D.19

3rd and 5th Grade Cohorts Not Eligible for Free or Reduced-Price School Lunch: Comparison of Significance Levels, Tests Grouped by Type

Grade 3 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*				*	*	*				*	*	*
TOWRE PDE	*	*	*	*	�	�				*	*	*	*	*	*	*	*	*
Word Identification	*	*	*	*	▓	▓	*	*	*									
TOWRE SWE	*	▓	*				*	▓	*									
AIMSweb	*	*	*	*	*	*	*	*	*				*	▓	▓			
Passage Comprehension	*	*	*				*	*	*									
GRADE																		

Grade 5 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*				*	*	*	*	*	*			
TOWRE PDE	*	*	*	*	*	*				*	*	*	*	▓	*			
Word Identification																		
TOWRE SWE													*	▓	▓			
AIMSweb				*	▓	▓				*	*	*				*	*	*
Passage Comprehension																		
GRADE																		

* Impact is statistically significant at the 0.05 level.

Note: "Unadj." is unadjusted p-values, "Bonf." referes to p-values adjusted using Bonferroni procedure, "BH" refers to p-values adjusted using Benjamini and Hochberg (1995) procedure.

Table D.20

3rd and 5th Grade Cohorts Not Eligible for Free or Reduced-Price School Lunch: Comparison of Significance Levels, All Tests Grouped Together

Grade 3 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	▓	*	*	▓	*				*	*	*	*	*	*	*	▓	*
TOWRE PDE	*	*	*	*	*	*				*	*	*	*	*	*	*	*	*
Word Identification	*	*	*	*	▓	▓	*	▓	*									
TOWRE SWE	*	▓	▓	*			*	▓	▓									
AIMSweb	*	*	*	*	▓	*	*	*	*				*	▓	▓			
Passage Comprehension GRADE	*	▓	*				*	*	*									

Grade 5 Cohort	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read			Wilson Reading			Corrective Reading		
	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH	Unadj.	Bonf.	BH
Word Attack	*	*	*	*	*	*				*	*	*	*	*	*			
TOWRE PDE	*	▓	▓	*	*	*				*	*	*	*	▓	▓			
Word Identification																		
TOWRE SWE													*	▓	▓			
AIMSweb				*	▓	▓				*	▓	*				*	*	*
Passage Comprehension GRADE																		

* Impact is statistically significant at the 0.05 level.
Note: "Unadj." is unadjusted p-values, "Bonf." referes to p-values adjusted using Bonferroni procedure, "BH" refers to p-values adjusted using Benjamini and Hochberg (1995) procedure.

APPENDIX E

INTERVENTION IMPACTS ON SPELLING AND CALCULATION

Table E.1

Impacts for 3rd and 5th Grade Cohorts

		All Interventions		Word-level Interventions		Failure Free Reading		Spell Read		Wilson Reading		Corrective Reading	
Grade 3 Cohort	Baseline	Control Gain	ABCD Impact	Control Gain	BCD Impact	Control Gain	A Impact	Control Gain	B Impact	Control Gain	C Impact	Control Gain	D Impact
WJ-III Spelling	88.5	2.4	1.3 *	2.9	0.9	0.8	2.6	2.3	0.9	3.1	0.9	3.3	0.8
WJ-III Calculation	95.4	1.9	4.3 *	2.4	3.6	0.6	6.6 *	5.3	0.4	-4.0	7.6	5.8	2.7
Sample Size	328												

		All Interventions		Word-level Interventions		Failure Free Reading		Spell Read		Wilson Reading		Corrective Reading	
Grade 5 Cohort	Baseline	Control Gain	ABCD Impact	Control Gain	BCD Impact	Control Gain	A Impact	Control Gain	B Impact	Control Gain	C Impact	Control Gain	D Impact
WJ-III Spelling	90.9	1.6	-0.5 #	1.4	-0.2	2.2	-1.5 * #	1.5	1.2	0.3	-0.2	2.5	-1.4
WJ-III Calculation	94.2	5.4	-1.5 * #	4.4	-0.8	8.3	-3.7 * #	2.2	0.7	4.8	-2.9 * #	6.2	-0.2
Sample Size	399												

* Impact is statistically significant at the 0.05 level.
# Impact is significantly different than the impact for the 3rd grade cohort at the 0.05 level.

Table E.2

Impacts for 3rd and 5th Grade Cohorts with Low Baseline Word Attack Scores

		All Interventions		Word-level Interventions		Failure Free Reading		Spell Read		Wilson Reading		Corrective Reading	
		Control	ABCD	Control	BCD	Control	A	Control	B	Control	C	Control	D
Grade 3 Cohort	Baseline	Gain	Impact	Gain	Impact	Gain	Impact	Gain	Impact	Gain	Impact	Gain	Impact
WJ-III Spelling	83.1	3.7	0.4	4.7	-0.3	0.7	2.5	4.5	-0.2	5.5	-0.2	4.2	-0.5
WJ-III Calculation	91.8	0.9	5.6 * #	0.2	7.3 * #	3.3	0.4	5.8	1.6 #	-7.4	11.9	2.1	8.5
Sample Size	169												

		All Interventions		Word-level Interventions		Failure Free Reading		Spell Read		Wilson Reading		Corrective Reading	
		Control	ABCD	Control	BCD	Control	A	Control	B	Control	C	Control	D
Grade 5 Cohort	Baseline	Gain	Impact	Gain	Impact	Gain	Impact	Gain	Impact	Gain	Impact	Gain	Impact
WJ-III Spelling	86.1	0.3	0.8	-0.4	1.6	2.3	-1.6	1.8	0.5	-1.1	0.9	-1.8	3.3 * #
WJ-III Calculation	92.4	5.3	-1.7	4.2	-0.1 #	8.3	-6.6 *	1.2	2.2	5.2	-2.1	6.3	-0.5
Sample Size	194												

* Impact is statistically significant at the 0.05 level.
# Impact is statistically different from the impact for all students in that grade at the 0.05 level.

Table E.3

Impacts for 3rd and 5th Grade Cohorts with High Baseline Word Attack Scores

Grade 3 Cohort	Baseline	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read		Wilson Reading			Corrective Reading	
		Control Gain	ABCD Impact		Control Gain	BCD Impact		Control Gain	A Impact		Control Gain	B Impact	Control Gain	C Impact		Control Gain	D Impact
WJ-III Spelling	94.3	1.5	1.9	*	1.4	1.3		1.8	3.6	*	0.7	1.0	1.0	1.3		2.5	1.5
WJ-III Calculation	99.1	5.9	0.3	#	7.6	-3.5	* #	0.7	11.9	*	4.6	-0.8	6.4	-5.2	#	11.7	-4.5 *
Sample Size	159																

Grade 5 Cohort	Baseline	All Interventions			Word-level Interventions			Failure Free Reading			Spell Read		Wilson Reading			Corrective Reading	
		Control Gain	ABCD Impact		Control Gain	BCD Impact		Control Gain	A Impact		Control Gain	B Impact	Control Gain	C Impact		Control Gain	D Impact
WJ-III Spelling	95.6	2.5	-1.5	*	2.3	-1.2		3.1	-2.4		1.3	2.1	1.8	-1.1		3.9	-4.7 *
WJ-III Calculation	96.0	5.1	-1.3		4.6	-1.7	#	6.8	-0.2		4.0	-1.6	3.8	-2.5		5.9	-1.0
Sample Size	205																

* Impact is statistically significant at the 0.05 level.
# Impact is statistically different from the impact for all students in that grade at the 0.05 level.

Table E.4

Impacts for 3rd and 5th Grade Cohorts with Low Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	Baseline	All Interventions Control Gain	ABCD Impact	Word-level Interventions Control Gain	BCD Impact	Failure Free Reading Control Gain	A Impact	Spell Read Control Gain	B Impact	Wilson Reading Control Gain	C Impact	Corrective Reading Control Gain	D Impact
WJ-III Spelling	87.8	3.9	-1.1	4.7	-1.6	1.5	0.4	5.9	-1.5	5.4	-1.7	2.7	-1.4 *
WJ-III Calculation	93.3	6.0	-0.5	6.8	-1.3	3.3	1.9	5.7	1.2	6.1	-2.9	8.7	-2.2 * #
Sample Size	146												

Grade 5 Cohort	Baseline	All Interventions Control Gain	ABCD Impact	Word-level Interventions Control Gain	BCD Impact	Failure Free Reading Control Gain	A Impact	Spell Read Control Gain	B Impact	Wilson Reading Control Gain	C Impact	Corrective Reading Control Gain	D Impact
WJ-III Spelling	89.2	2.1	-0.7	1.7	-0.1	3.3	-2.6 *	3.0	0.7	2.1	-1.1	0.2	0.2
WJ-III Calculation	92.2	3.3	0.5	2.0	1.4	7.0	-2.2	-0.6	4.0 * #	1.9	0.0	4.9	0.3
Sample Size	194												

* Impact is statistically significant at the 0.05 level.
# Impact is statistically different from the impact for all students in that grade at the 0.05 level.

Table E.5

Impacts for 3rd and 5th Grade Cohorts with High Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	Baseline	All Interventions		Word-level Interventions		Failure Free Reading		Spell Read		Wilson Reading		Corrective Reading	
		Control Gain	ABCD Impact	Control Gain	BCD Impact	Control Gain	A Impact	Control Gain	B Impact	Control Gain	C Impact	Control Gain	D Impact
WJ-III Spelling	89.1	2.4	1.7 *	2.8	1.1	1.1	3.4 *	1.0	1.0	3.9	0.4	3.4	1.9
WJ-III Calculation	97.0	1.3	4.3 *	1.4	3.7	1.0	6.1	4.0	0.0	-7.8	10.1	7.9	1.1 #
Sample Size	182												

Grade 5 Cohort	Baseline	All Interventions		Word-level Interventions		Failure Free Reading		Spell Read		Wilson Reading		Corrective Reading	
		Control Gain	ABCD Impact	Control Gain	BCD Impact	Control Gain	A Impact	Control Gain	B Impact	Control Gain	C Impact	Control Gain	D Impact
WJ-III Spelling	92.4	1.5	-1.0	1.5	-1.0	1.4	-0.9	0.8	0.6	0.9	-1.8	2.7	-1.9
WJ-III Calculation	95.9	6.4	-2.5 *	5.7	-1.9	8.5	-4.3	5.6	-3.2 #	6.1	-3.3 *	5.4	1.0
Sample Size	205												

* Impact is statistically significant at the 0.05 level.
# Impact is statistically different from the impact for all students in that grade at the 0.05 level.

Table E.6

Impacts for 3rd and 5th Grade Cohorts with Low Baseline Word Attack and Low Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	Baseline	All Interventions		Word-level Interventions		Failure Free Reading		Spell Read		Wilson Reading		Corrective Reading	
		Control Gain	ABCD Impact	Control Gain	BCD Impact	Control Gain	A Impact	Control Gain	B Impact	Control Gain	C Impact	Control Gain	D Impact
WJ-III Spelling	83.0	6.2	-2.4 * #	6.7	-2.9 * #	4.6	-1.1	8.0	-2.4	8.2	-1.2	3.9	-5.1
WJ-III Calculation	89.6	3.5	2.8	3.3	4.2	4.1	-1.4 #	10.5	-4.4	-2.3	10.9	1.9	6.3
Sample Size	80												

Grade 5 Cohort	Baseline	All Interventions		Word-level Interventions		Failure Free Reading		Spell Read		Wilson Reading		Corrective Reading	
		Control Gain	ABCD Impact	Control Gain	BCD Impact	Control Gain	A Impact	Control Gain	B Impact	Control Gain	C Impact	Control Gain	D Impact
WJ-III Spelling	84.3	2.7	-2.0 *	2.7	-2.0	2.6	-2.2	4.1	-2.2 #	2.8	-3.1	1.2	-0.5 *
WJ-III Calculation	89.9	4.7	-1.0	4.2	0.0	5.9	-3.9	-2.9	5.4 * #	7.3	-3.6	8.3	-1.9
Sample Size	106												

* Impact is statistically significant at the 0.05 level.
# Impact is statistically different from the impact for all students in that grade at the 0.05 level.

Table E.7

Impacts for 3rd and 5th Grade Cohorts With Low Baseline Word Attack and High Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	Baseline	All Interventions		Word-level Interventions		Failure Free Reading		Spell Read		Wilson Reading		Corrective Reading	
		Control Gain	ABCD Impact	Control Gain	BCD Impact	Control Gain	A Impact	Control Gain	B Impact	Control Gain	C Impact	Control Gain	D Impact
WJ-III Spelling	83.1	2.8	1.9	3.5	1.5	0.8	3.1 *	4.2	-0.5	4.6	0.7	1.7	4.5
WJ-III Calculation	93.9	2.0	3.1	1.0	4.6 #	4.9	-1.3	3.4	3.2	-11.0	10.8	10.7	-0.3
Sample Size	89												

Grade 5 Cohort	Baseline	All Interventions		Word-level Interventions		Failure Free Reading		Spell Read		Wilson Reading		Corrective Reading	
		Control Gain	ABCD Impact	Control Gain	BCD Impact	Control Gain	A Impact	Control Gain	B Impact	Control Gain	C Impact	Control Gain	D Impact
WJ-III Spelling	87.9	-1.2	1.4 #	-1.8	2.1 #	0.8	-0.9	0.6	1.6	-1.7	-0.2	-4.4	5.0 * #
WJ-III Calculation	94.8	5.5	-2.4	4.2	-0.7	9.4	-7.8 * #	4.0	0.5	5.3	-3.5	3.4	1.0
Sample Size	88												

* Impact is statistically significant at the 0.05 level.
# Impact is statistically different from the impact for all students in that grade at the 0.05 level.

Table E.8

Impacts for 3rd and 5th Grade Cohorts with High Baseline Word Attack and High Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	Baseline	All Interventions		Word-level Interventions		Failure Free Reading		Spell Read		Wilson Reading		Corrective Reading	
		Control Gain	ABCD Impact	Control Gain	BCD Impact	Control Gain	A Impact	Control Gain	B Impact	Control Gain	C Impact	Control Gain	D Impact
WJ-III Spelling	94.8	3.3	-0.1	3.4	-1.0	2.8	2.5	-0.9	1.7	5.9	-2.7	5.3	-1.9
WJ-III Calculation	100.0	3.0	2.7	4.5	-0.6 #	-1.6	12.7 * #	6.2	-4.5 #	-0.6	3.8	7.9	-1.0
Sample Size	93												

Grade 5 Cohort	Baseline	All Interventions		Word-level Interventions		Failure Free Reading		Spell Read		Wilson Reading		Corrective Reading	
		Control Gain	ABCD Impact	Control Gain	BCD Impact	Control Gain	A Impact	Control Gain	B Impact	Control Gain	C Impact	Control Gain	D Impact
WJ-III Spelling	96.0	2.7	-2.5 * #	2.8	-2.7 * #	2.4	-1.9	0.6	0.0	2.6	-2.1	5.2	-5.9 * #
WJ-III Calculation	96.8	5.9	-2.2	5.6	-2.6	6.8	-1.0	6.9	-7.3 #	4.6	-1.3	5.2	0.8
Sample Size	117												

* Impact is statistically significant at the 0.05 level.
# Impact is statistically different from the impact for all students in that grade at the 0.05 level.

Table E.9

Impacts for 3rd and 5th Grade Cohorts Eligible for Free or Reduced Price School Lunch

Grade 3 Cohort	Baseline	All Interventions		Word-level Interventions		Failure Free Reading		Spell Read		Wilson Reading		Corrective Reading	
		Control Gain	ABCD Impact	Control Gain	BCD Impact	Control Gain	A Impact	Control Gain	B Impact	Control Gain	C Impact	Control Gain	D Impact
WJ-III Spelling	87.9	1.8	2.4	3.1	1.3	-2.0	5.8 *	3.3	0.0	3.4	1.5	2.6	2.4
WJ-III Calculation	95.1	2.8	3.1	3.8	2.2	-0.1	5.7	1.9	3.3	2.6	0.9	6.9	2.4
Sample Size	189												

Grade 5 Cohort	Baseline	All Interventions		Word-level Interventions		Failure Free Reading		Spell Read		Wilson Reading		Corrective Reading	
		Control Gain	ABCD Impact	Control Gain	BCD Impact	Control Gain	A Impact	Control Gain	B Impact	Control Gain	C Impact	Control Gain	D Impact
WJ-III Spelling	89.3	1.3	-1.4	1.5	-1.3	0.8	-1.8	2.6	-1.8 #	1.3	-1.0	0.5	-1.0
WJ-III Calculation	92.8	5.9	-4.4 *	4.4	-3.6	10.4	-6.9 *	3.9	-3.6	2.9	-3.1	6.4	-4.0 * #
Sample Size	219												

* Impact is statistically significant at the 0.05 level.
# Impact is statistically different from the impact for all students in that grade at the 0.05 level.

Table E.10

Impacts for 3rd and 5th Grade Cohorts Not Eligible for Free or Reduced Price School Lunch

Grade 3 Cohort	Baseline	All Interventions		Word-level Interventions		Failure Free Reading		Spell Read		Wilson Reading		Corrective Reading	
		Control Gain	ABCD Impact	Control Gain	BCD Impact	Control Gain	A Impact	Control Gain	B Impact	Control Gain	C Impact	Control Gain	D Impact
WJ-III Spelling	89.4	1.7	2.3	1.0	2.9	3.9	0.2	1.3	2.5 *	-1.9	5.5	3.5	0.9
WJ-III Calculation	95.8	-3.3	9.0	-4.4	9.4	0.0	7.8	5.5	-0.7	-22.1	25.5	3.5	3.5
Sample Size	139												

Grade 5 Cohort	Baseline	All Interventions		Word-level Interventions		Failure Free Reading		Spell Read		Wilson Reading		Corrective Reading	
		Control Gain	ABCD Impact	Control Gain	BCD Impact	Control Gain	A Impact	Control Gain	B Impact	Control Gain	C Impact	Control Gain	D Impact
WJ-III Spelling	92.7	2.2	-1.5 *	2.0	-1.5	2.8	-1.5	-0.2	3.3 * #	1.0	-3.5	5.3	-4.3 *
WJ-III Calculation	95.8	5.8	-1.1	5.4	-0.4	6.8	-3.2	3.4	1.3	7.2	-4.9 *	5.7	2.2 #
Sample Size	180												

* Impact is statistically significant at the 0.05 level.
# Impact is statistically different from the impact for all students in that grade at the 0.05 level.

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*
WJ-III Spelling	3.85	0.05	1.87	0.17	2.96	0.08	0.46	0.50	1.12	0.29	0.93	0.50
WJ-III Calculation	5.34	0.02	2.58	0.10	4.51	0.03	0.02	0.50	2.19	0.13	0.75	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*
WJ-III Spelling	0.60	0.50	0.04	0.50	4.04	0.04	0.80	0.50	0.05	0.50	0.92	0.50
WJ-III Calculation	4.02	0.04	0.98	0.50	4.58	0.03	0.23	0.50	9.92	0.00	0.02	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
"*p*" is the *p*-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table E.12

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts with Low Baseline Word Attack Scores

	All Interventions		Word-level Interventions		Failure Free Reading		Spell Read		Wilson Reading		Corrective Reading	
	ABCD		BCD		A		B		C		D	
Grade 3 Cohort	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*
WJ-III Spelling	0.31	0.50	0.16	0.50	1.82	0.17	0.03	0.50	0.02	0.50	0.31	0.50
WJ-III Calculation	4.42	0.03	4.67	0.03	0.01	0.50	0.21	0.50	3.04	0.08	1.48	0.22

	All Interventions		Word-level Interventions		Failure Free Reading		Spell Read		Wilson Reading		Corrective Reading	
	ABCD		BCD		A		B		C		D	
Grade 5 Cohort	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*
WJ-III Spelling	0.86	0.50	2.50	0.11	2.10	0.14	0.12	0.50	0.35	0.50	5.47	0.02
WJ-III Calculation	1.58	0.21	0.01	0.50	9.76	0.00	1.61	0.20	1.31	0.25	0.02	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
    "*p*" is the *p*-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table E.13

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts with High Baseline Word Attack Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*
WJ-III Spelling	7.01	0.01	2.34	0.12	7.46	0.01	0.41	0.50	1.68	0.19	1.12	0.29
WJ-III Calculation	0.06	0.50	4.27	0.04	21.76	0.00	0.14	0.50	2.93	0.08	5.07	0.02

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*
WJ-III Spelling	3.89	0.05	2.11	0.14	2.84	0.09	1.20	0.27	1.17	0.28	11.42	0.00
WJ-III Calculation	1.05	0.31	1.34	0.25	0.01	0.50	0.24	0.50	3.01	0.08	0.69	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
   "*p*" is the *p*-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table E.14

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts with Low Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*
WJ-III Spelling	0.98	0.50	1.91	0.16	0.02	0.50	0.90	0.50	0.76	0.50	0.31	0.50
WJ-III Calculation	0.05	0.50	0.30	0.50	0.27	0.50	0.13	0.50	0.84	0.50	0.33	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*
WJ-III Spelling	1.04	0.31	0.01	0.50	10.10	0.00	0.41	0.50	0.65	0.50	0.03	0.50
WJ-III Calculation	0.16	0.50	0.75	0.50	1.33	0.25	7.20	0.01	0.00	0.50	0.01	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
    "*p*" is the *p*-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table E.15

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts with High Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
WJ-III Spelling	5.22	0.02	1.50	0.22	5.77	0.02	0.39	0.50	0.08	0.50	1.23	0.27
WJ-III Calculation	3.84	0.05	2.19	0.13	2.90	0.08	0.00	0.50	2.52	0.11	0.17	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
WJ-III Spelling	1.03	0.31	0.65	0.50	0.75	0.50	0.10	0.50	0.79	0.50	1.15	0.28
WJ-III Calculation	5.06	0.02	2.69	0.10	2.98	0.08	1.77	0.18	7.27	0.01	0.30	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
"p"is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Title: Table E.16
Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts with Low Baseline Word Attack and Low Screening Peabody Picture Vocabulary Test Scores

Headers span: All Interventions (ABCD), Word-level Interventions (BCD), Failure Free Reading (A), Spell Read (B), Wilson Reading (C), Corrective Reading (D). Each has Test and p columns.

Grade 3 Cohort:
WJ-III Spelling: 3.81, 0.05, 6.00, 0.01, 0.07, 0.50, 1.35, 0.24, 0.23, 0.50, 1.98, 0.16
WJ-III Calculation: 0.63, 0.50, 0.92, 0.50, 0.08, 0.50, 0.65, 0.50, 1.39, 0.24, 0.53, 0.50

Grade 5 Cohort:
WJ-III Spelling: 4.51, 0.03, 2.10, 0.14, 1.62, 0.20, 2.78, 0.09, 2.58, 0.10, 0.04, 0.50
WJ-III Calculation: 0.30, 0.50, 0.00, 0.50, 2.00, 0.15, 9.29, 0.00, 0.94, 0.50, 0.14, 0.50Table E.16

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts with Low Baseline Word Attack and Low Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*
WJ-III Spelling	3.81	0.05	6.00	0.01	0.07	0.50	1.35	0.24	0.23	0.50	1.98	0.16
WJ-III Calculation	0.63	0.50	0.92	0.50	0.08	0.50	0.65	0.50	1.39	0.24	0.53	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*
WJ-III Spelling	4.51	0.03	2.10	0.14	1.62	0.20	2.78	0.09	2.58	0.10	0.04	0.50
WJ-III Calculation	0.30	0.50	0.00	0.50	2.00	0.15	9.29	0.00	0.94	0.50	0.14	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
"*p*" is the *p*-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table E.17

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts With Low Baseline Word Attack and High Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
WJ-III Spelling	1.83	0.17	0.71	0.50	4.36	0.03	0.05	0.50	0.05	0.50	1.38	0.24
WJ-III Calculation	1.35	0.24	2.18	0.14	0.11	0.50	1.64	0.20	2.15	0.14	0.00	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
WJ-III Spelling	2.23	0.13	3.47	0.06	0.38	0.50	0.99	0.50	0.02	0.50	5.02	0.02
WJ-III Calculation	3.63	0.05	0.17	0.50	13.36	0.00	0.08	0.50	1.64	0.20	0.06	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
  "p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table E.18

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts with High Baseline Word Attack and High Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
WJ-III Spelling	0.02	0.50	1.11	0.29	2.16	0.14	1.70	0.19	1.44	0.23	1.89	0.17
WJ-III Calculation	1.55	0.21	0.05	0.50	13.18	0.00	2.26	0.13	0.57	0.50	0.09	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
WJ-III Spelling	7.96	0.01	6.05	0.01	1.32	0.25	0.00	0.50	0.73	0.50	13.86	0.00
WJ-III Calculation	1.64	0.20	1.64	0.20	0.10	0.50	2.41	0.12	0.27	0.50	0.30	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
"p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table E.19

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts Eligible for Free or Reduced Price School Lunch

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
WJ-III Spelling	3.60	0.05	0.98	0.50	5.15	0.02	0.00	0.50	1.04	0.31	1.34	0.25
WJ-III Calculation	3.05	0.08	1.71	0.19	2.78	0.09	1.10	0.29	0.12	0.50	1.56	0.21

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
WJ-III Spelling	2.41	0.12	1.42	0.23	2.10	0.14	0.85	0.50	0.58	0.50	0.56	0.50
WJ-III Calculation	5.49	0.02	2.91	0.08	8.59	0.00	1.78	0.18	1.19	0.27	3.20	0.07

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
   "p"is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table E.20

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts Not Eligible for Free or Reduced Price School Lunch

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
WJ-III Spelling	2.57	0.11	2.50	0.11	0.02	0.50	3.72	0.05	1.19	0.27	0.26	0.50
WJ-III Calculation	3.33	0.06	2.33	0.12	1.51	0.22	0.05	0.50	2.48	0.11	0.23	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
WJ-III Spelling	3.99	0.04	2.72	0.10	1.99	0.15	10.85	0.00	2.65	0.10	6.13	0.01
WJ-III Calculation	1.42	0.23	0.17	0.50	1.81	0.17	0.67	0.50	17.15	0.00	1.27	0.26

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
"p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table E.21

Tests for Differences in Impacts for 3rd and 5th Grade Cohorts

	All Interventions		Word-level Interventions		Failure Free Reading		Spell Read		Wilson Reading		Corrective Reading	
	ABCD		BCD		A		B		C		D	
	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*
WJ-III Spelling	3.84	0.05	1.05	0.31	5.87	0.02	0.01	0.91	0.70	0.40	1.73	0.19
WJ-III Calculation	8.37	0.00	3.42	0.06	8.39	0.00	0.01	0.92	4.02	0.05	0.70	0.40

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
"*p*" is the *p*-value produced by HLM6 software.

Table E.22

Tests for Differences in Impacts for 3rd and 5th Grade Cohorts with Low Baseline Word Attack Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
WJ-III Spelling	2.60	0.10	2.81	0.09	0.15	0.50	0.64	0.50	1.07	0.30	1.61	0.20
WJ-III Calculation	3.15	0.07	7.82	0.01	11.38	0.00	0.45	0.50	4.90	0.03	2.79	0.09

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
WJ-III Spelling	3.17	0.07	3.52	0.06	0.16	0.50	0.41	0.50	0.96	0.50	19.31	0.00
WJ-III Calculation	0.03	0.50	0.35	0.50	4.79	0.03	0.85	0.50	0.02	0.50	0.01	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
   "p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table E.23

Tests for Differences in Impacts for 3rd and 5th Grade Cohorts with Low Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
WJ-III Spelling	3.56	0.06	2.27	0.13	0.81	0.50	1.19	0.27	0.44	0.50	0.74	0.50
WJ-III Calculation	3.42	0.06	3.25	0.07	1.43	0.23	0.10	0.50	6.62	0.01	1.80	0.18

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
WJ-III Spelling	0.06	0.50	0.37	0.50	0.94	0.50	0.00	0.50	0.09	0.50	0.94	0.50
WJ-III Calculation	2.59	0.10	2.74	0.09	0.40	0.50	5.61	0.02	1.20	0.27	0.03	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
"p"is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table E.24

Tests for Differences in Impacts for 3rd and 5th Grade Cohorts with Low Baseline Word Attack and Low Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
WJ-III Spelling	10.55	0.00	9.19	0.00	0.85	0.50	1.79	0.18	0.96	0.50	3.09	0.08
WJ-III Calculation	0.28	0.50	0.13	0.50	5.32	0.02	1.65	0.20	0.88	0.50	0.26	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
WJ-III Spelling	1.83	0.17	1.22	0.27	0.26	0.50	4.90	0.03	0.88	0.50	0.21	0.50
WJ-III Calculation	0.09	0.50	0.14	0.50	0.03	0.50	4.10	0.04	0.05	0.50	0.08	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
   "p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table E.25

Tests for Differences in Impacts for 3rd and 5th Grade Cohorts With Low Baseline Word Attack and High Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
WJ-III Spelling	0.36	0.50	0.14	0.50	0.50	0.50	0.30	0.50	0.00	0.50	0.63	0.50
WJ-III Calculation	1.27	0.26	4.32	0.04	3.68	0.05	1.23	0.27	3.52	0.06	0.64	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
WJ-III Spelling	5.92	0.01	5.80	0.02	0.59	0.50	0.37	0.50	0.00	0.50	10.97	0.00
WJ-III Calculation	0.71	0.50	0.00	0.50	5.23	0.02	0.10	0.50	0.02	0.50	0.10	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
　　"p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table E.26

Tests for Differences in Impacts for 3rd and 5th Grade Cohorts with High Baseline Word Attack and High Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
WJ-III Spelling	1.85	0.17	2.82	0.09	0.00	0.50	0.81	0.50	3.44	0.06	1.99	0.15
WJ-III Calculation	0.84	0.50	4.04	0.04	6.62	0.01	4.62	0.03	0.33	0.50	1.80	0.18

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
WJ-III Spelling	11.73	0.00	13.29	0.00	0.03	0.50	0.55	0.50	2.47	0.11	10.30	0.00
WJ-III Calculation	0.19	0.50	1.06	0.31	0.98	0.50	4.61	0.03	0.03	0.50	0.30	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
   "p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table E.27

Tests for Differences in Impacts for 3rd and 5th Grade Cohorts Eligible for Free or Reduced Price School Lunch

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
WJ-III Spelling	0.01	0.50	0.48	0.50	3.19	0.07	1.00	0.32	0.57	0.50	0.20	0.50
WJ-III Calculation	1.38	0.24	1.25	0.26	0.15	0.50	1.58	0.21	2.15	0.14	0.02	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
WJ-III Spelling	0.00	0.50	0.02	0.50	0.02	0.50	4.60	0.03	1.17	0.28	2.07	0.15
WJ-III Calculation	2.40	0.12	1.65	0.20	1.23	0.27	1.84	0.17	0.41	0.50	4.60	0.03

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
"p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

APPENDIX F

INSTRUCTIONAL GROUP CLUSTERING

For treatment students, the interventions were administered in instructional groups of three students each. Control students do not have instructional groups; however, for purposes of this analysis, students in the control group within each school unit and grade level are treated as if they came from a single instructional group. Typically, for treatment students, the instructional group consists of three students in the same grade, although 5 of the 146 treatment instructional groups had a mixture of third and fifth graders, either because of scheduling difficulties or because a student fit better with students from the other grade in terms of reading ability. The 71 control "instructional groups" ranged in size from 1 to 18 students, with a median size of 3 students.

The instructional groups may introduce clustering among students for which our two-level model does not account; such clustering may affect the standard errors of our impact estimates. To assess the effect of this instructional group clustering on the standard errors of our impact estimates, we expanded our two-level model to a three-level model that includes a level for instructional groups. Specifically, the three-level model is:

Level One: Student i in instructional group j and School-unit k

$$y_{ijk} = \pi_{0jk} + \pi_{1jk} y^{*}_{oijk} + e_{ijk}$$

Level Two: Instructional group j and school-unit k

$$\pi_{0jk} = \beta_{00k} + \beta_{01k} T_{jk} + r_{0jk}$$
$$\pi_{1jk} = \beta_{10k} + r_{1jk}$$

Level Three: School-unit k

$$\beta_{00k} = \gamma_{000} + \gamma_{001} A_k + \gamma_{002} B_k + \gamma_{003} C_k + \sum_{l=1}^{3} \xi_{00l} P_{lk} + \mu_{00k}$$

$$\beta_{01k} = \gamma_{010} + \gamma_{011} A_k + \gamma_{012} B_k + \gamma_{013} C_k + \sum_{l=1}^{3} \xi_{01l} P_{lk} + \mu_{01k}$$

$$\beta_{10k} = \gamma_{100} + \gamma_{101} A_k + \gamma_{102} B_k + \gamma_{103} C_k + \sum_{l=1}^{3} \xi_{10l} P_{lk} + \mu_{10k}.$$

To assess the sensitivity of the impact estimates and their standard errors to the incorporation of instructional group clustering, we examine the impact estimates obtained from the two- and three-level models and the ratio of the variance of the impact estimate from the three-level model to the variance from the two-level model (see Table F.1). We find that impact estimates from the two- and three-level models are generally similar. We also find that the two-level model generally yields slightly higher variances than the three-level model, with a few exceptions, including some impacts that are not estimated very precisely and thus have relatively large variances. For such impacts, the ratio of variances is practically unimportant because it corresponds to an impact estimated with high sampling variability—in other words, a large standard error relative to the impact estimate, and thus the impact estimate

cannot be statistically distinguished from zero. For example, on the Passage Comprehension test for the third grade cohort, the word-level intervention impact (BCD) is measured with low precision in both the two-level model and the three-level model, thus making the small ratio of variances practically unimportant. Overall, there is generally a gain in precision when using the three-level model; however, the differences do not change the substantive conclusions, and would change the significance level of an impact only in a case with borderline significance. We therefore base our conclusions on impact estimates derived from the two-level model.

Table F.1

Impact for 3rd and 5th Grade Cohorts for Three-Level Model

Grade 3 Cohort	All Interventions	Word-level Interventions	Failure Free Reading	Spell Read	Wilson Reading	Corrective Reading
	ABCD	BCD	A	B	C	D
	Impact	Impact	Impact	Impact	Impact	Impact
Word Attack	4.58 *	4.85 *	3.79 *	5.07 *	5.17 *	4.30 *
TOWRE PDE	3.46 *	4.30 *	0.95	5.10 *	3.90 *	3.90
Word Identification	2.17 *	2.47 *	1.26	1.15	4.07 *	2.18
TOWRE SWE	2.19 *	2.30 *	1.86 *	0.82	3.19 *	2.88
AIMSweb	6.30 *	6.22	6.53 *	6.70 *	4.29	7.68
Passage Comprehension	1.95 *	1.35	3.75 *	-0.52	2.99	1.59
GRADE	0.79	0.07	2.95	1.30	-0.48	-0.61

Grade 3 Cohort	All Interventions	Word-level Interventions	Failure Free Reading	Spell Read	Wilson Reading	Corrective Reading
	ABCD	BCD	A	B	C	D
	Impact	Impact	Impact	Impact	Impact	Impact
Word Attack	2.39 *	3.49 *	-0.90	3.53 *	6.50 *	0.44
TOWRE PDE	1.83 *	2.38 *	0.17	3.47 *	1.45	2.21
Word Identification	-0.75	-0.79	-0.62	-1.27	0.16	-1.27
TOWRE SWE	1.71 *	1.68	1.78	3.80 *	0.59	0.65
AIMSweb	-2.85	-2.66	-3.42	-2.69	-1.61	-3.67
Passage Comprehension	-1.06	-0.76	-1.97	-0.77	0.26	-1.77
GRADE	0.51	0.83	-0.45	-0.45	3.55	-0.63

* Impact is statistically significant at the 0.05 level; shaded cells indicate a change in significance relative to the two-level model.

APPENDIX G

BASELINE CHARACTERISTICS OF THE PSSA SAMPLE

Table G.1

Baseline Characteristics of the Four Intervention Groups and the Control Group,
Analysis Sample: 3rd Grade (PSSA Sample)

Baseline Means	Failure Free Reading		Spell Read		Wilson Reading		Corrective Reading	
	Treat.	Cont.	Treat.	Cont.	Treat.	Cont.	Treat.	Cont.
Student Characteristics								
Age	8.5	8.7	8.8	8.7	8.7	8.6	8.7	8.6
Male (%)	53	58	73	59	39	21	56	44
Hispanic (%)	a	a	a	a	a	a	a	a
Race--White (%)	77	83	65	68	55	68	74	82
Race--African American (%)	23	17	35	32	45	32	26	18
Race--Other (%)	a	a	a	a	a	a	a	a
Family income less than $30,000 (%)	43	39	57	49	48	63	41	60
Family income between $30,000 and $60,000 (%)	46	45	20	24	32	37	41	26
Family income over $60,000 (%)	11	15	23	27	a	a *	18	14
Eligible for free or reduced price lunch (%)	47	47	46	36	36	70 *	42	49
Has any learning or other disability (%)	40	48	35	25	34	19	30	32
Mother has bachelor's degree or higher (%)	14	9 *	13	15	a	a	19	11
Screening Tests								
TOWRE Sight Word Efficiency	84.2	81.9	85.4	84.7	86.8	83.9	85.6	83.7
TOWRE Phonemic Decoding Efficiency	84.4	85.4	85.7	85.0	86.1	85.8	85.9	87.5
Peabody Picture Vocabulary Test--Revised	93.9	95.2	95.5	97.8	90.4	89.8	97.4	99.9
Baseline Tests								
WRM Word Identification	88.9	87.4	89.5	87.2 *	90.6	89.2	89.7	89.1
TOWRE Phonemic Decoding Efficiency	84.5	84.2	86.2	84.6	87.0	86.3	87.3	87.0
WRM Word Attack	90.2	89.1	93.8	91.4	94.7	94.2	93.9	95.8
TOWRE Sight Word Efficiency	87.3	84.6	89.3	86.6 *	89.0	83.3 *	87.0	86.5
AIMSWeb (Raw score)	38.4	33.5	46.8	41.4	49.3	38.0 *	43.5	41.5
WRM Passage Comprehension	90.9	88.9	95.2	89.9 *	93.8	92.0	94.2	92.9
GRADE	86.4	85.0	87.8	83.9	88.6	85.1	89.7	88.2
Woodcock Johnson Spelling	90.0	86.8	89.4	89.0	89.3	86.7	90.3	87.5
Woodcock Johnson Calculation	93.2	97.1 *	99.3	95.1 *	96.9	91.1	96.8	98.3
Sample Size	53	36	56	36	53	18	43	34

Note: Weights used to account for differential randomization probabilities and nonresponse.

Note: All test scores are shown as standard scores, unless otherwise indicated. All standard scores have mean 100 and standard deviation 15, except for CTOPP and Clinical Evaluation of Language Fundamentals-IV, which have mean 10 and standard deviation 3.

* Difference between treatment and control groups is statistically significant at the 0.05 level.

Difference across the four interventions (with treatment and control groups pooled within each intervention) is statistically significant at the 0.05 level.

a Values suppressed to protect student confidentiality.

G-3

Table G.2

Baseline Characteristics of the Four Intervention Groups and the Control Group,
Analysis Sample: 5th Grade (PSSA Sample)

Baseline Means	Failure Free Reading		Spell Read		Wilson Reading			Corrective Reading		
	Treat.	Cont.	Treat.	Cont.	Treat.	Cont.		Treat.	Cont.	
Student Characteristics										
Age	10.6	10.6	10.9	10.7	10.8	10.5	*	10.7	10.6	
Male (%)	53	51	55	58	54	66		49	59	
Hispanic (%)	a	a	a	a	a	a		a	a	
Race--White (%)	78	83	74	67	55	59		83	86	
Race--African American (%)	22	17	26	33	45	41		18	14	
Race--Other (%)	0	0	0	0	0	0		0	0	
Family income less than $30,000 (%)	41	50	51	54	73	47	*	34	52	
Family income between $30,000 and $60,000 (%)	43	33	39	34	23	36		47	36	
Family income over $60,000 (%)	16	17	11	12	4	18	*	19	12	
Eligible for free or reduced price lunch (%)	42	46	52	38	* 56	42	*	46	46	
Has any learning or other disability (%)	27	36	27	30	32	30		28	29	
Mother has bachelor's degree or higher (%)	12	17	5	9	0	17		13	24	*
Screening Tests										
TOWRE Sight Word Efficiency	84.1	85.3	84.1	85.3	83.5	84.8		82.9	83.4	
TOWRE Phonemic Decoding Efficiency	81.7	79.8	* 78.6	79.7	82.7	82.2		80.0	80.5	
Peabody Picture Vocabulary Test--Revised	94.8	95.1	92.0	92.4	91.6	100.0	*	95.2	97.3	
Baseline Tests										
WRM Word Identification	90.4	89.0	87.4	87.6	87.9	90.0		87.6	89.3	
TOWRE Phonemic Decoding Efficiency	82.1	81.8	78.0	79.9	82.9	81.2		80.6	81.3	
WRM Word Attack	93.5	92.7	91.0	92.2	93.3	94.4		93.6	93.7	
TOWRE Sight Word Efficiency	84.2	85.7	83.3	85.5	84.2	84.6		83.4	82.9	
AIMSWeb (Raw score)	79.0	75.6	80.2	79.3	74.7	80.2	*	75.0	71.6	
WRM Passage Comprehension	92.4	92.2	91.6	92.8	90.5	96.4	*	92.0	93.3	
GRADE	91.4	92.2	89.9	89.1	92.0	95.2	*	88.1	91.5	
Woodcock Johnson Spelling	93.8	92.2	90.1	91.3	91.1	92.3		88.3	86.9	
Woodcock Johnson Calculation	94.0	93.5	94.9	94.5	93.6	95.2		94.5	94.5	
Sample Size	62	66	56	46	54	38		55	31	

Note: Weights used to account for differential randomization probabilities and nonresponse.

Note: All test scores are shown as standard scores, unless otherwise indicated. All standard scores have mean 100 and standard deviation 15, except for
CTOPP and Clinical Evaluation of Language Fundamentals-IV, which have mean 10 and standard deviation 3.

* Difference between treatment and control groups is statistically significant at the 0.05 level.

Difference across the four interventions (with treatment and control groups pooled within each intervention) is statistically significant at the 0.05 level.

a Values suppressed to protect student confidentiality.

Baseline Characteristics of Full Sample and Three Word-level Interventions, by Treatment Status,
Analysis Sample: 3rd Grade (PSSA Sample)

Baseline Means	All Interventions			Word-level Interventions		
	Treatment	Control		Treatment	Control	
Student Characteristics						
Age	8.7	8.6		8.7	8.6	
Male (%)	56	47		57	42	*
Hispanic (%)	2	2		a	a	
Race--White (%)	68	75	*	65	72	
Race--African American (%)	32	25	*	35	28	
Race--Other (%)	a	a		a	a	
Family income less than $30,000 (%)	48	52		50	57	
Family income between $30,000 and $60,000 (%)	34	34		30	29	
Family income over $60,000 (%)	18	14		20	14	
Eligible for free or reduced price lunch (%)	43	50		42	51	
Has any learning or other disability (%)	35	31		33	25	
Mother has bachelor's degree or higher (%)	13	10		12	11	
Screening Tests						
TOWRE Sight Word Efficiency	85.5	83.5		85.9	84.1	
TOWRE Phonemic Decoding Efficiency	85.5	85.9		85.9	86.1	
Peabody Picture Vocabulary Test--Revised	94.3	95.7		94.4	95.8	
Baseline Tests						
WRM Word Identification	89.6	88.2	*	89.9	88.4	
TOWRE Phonemic Decoding Efficiency	86.2	85.4		86.8	85.9	
WRM Word Attack	93.1	92.4		94.1	93.7	
TOWRE Sight Word Efficiency	88.2	85.2	*	88.5	85.5	*
AIMSWeb (Raw score)	44.4	38.5	*	46.6	40.3	*
WRM Passage Comprehension	93.5	90.8		94.4	91.5	
GRADE	88.0	85.5	*	88.6	85.6	*
Woodcock Johnson Spelling	89.8	87.5		89.7	87.8	
Woodcock Johnson Calculation	96.5	95.4		97.7	94.8	*
Sample Size	205	124		152	88	

Note: Weights used to account for differential randomization probabilities and nonresponse.

Note: All test scores are shown as standard scores, unless otherwise indicated. All standard scores have mean 100 and standard deviation 15, except for CTOPP and Clinical Evaluation of Language Fundamentals-IV, which have mean 10 and standard deviation 3.

* Difference between treatment and control groups is statistically significant at the 0.05 level.

a Values suppressed to protect student confidentiality.

Table G.4

Baseline Characteristics of Full Sample and Three Word-level Interventions, by Treatment Status,
Analysis Sample: 5th Grade (PSSA Sample)

Baseline Means	All Interventions			Word-level Interventions		
	Treatment	Control		Treatment	Control	
Student Characteristics						
Age	10.7	10.6	*	10.8	10.6	*
Male (%)	53	58		53	61	
Hispanic (%)	a	a		a	a	
Race--White (%)	73	74		71	71	
Race--African American (%)	27	26		29	29	
Race--Other (%)	a	a		a	a	
Family income less than $30,000 (%)	49	51		51	51	
Family income between $30,000 and $60,000 (%)	39	35		37	35	
Family income over $60,000 (%)	13	14		12	14	
Eligible for free or reduced price lunch (%)	49	43		51	42	*
Has any learning or other disability (%)	28	31		29	30	
Mother has bachelor's degree or higher (%)	8	17	*	6	17	*
Screening Tests						
TOWRE Sight Word Efficiency	83.7	84.7		83.5	84.5	
TOWRE Phonemic Decoding Efficiency	80.7	80.5		80.4	80.8	
Peabody Picture Vocabulary Test--Revised	93.4	96.1	*	93.0	96.5	*
Baseline Tests						
WRM Word Identification	88.3	89.0		87.6	89.0	
TOWRE Phonemic Decoding Efficiency	80.8	81.1		80.4	80.8	
WRM Word Attack	92.9	93.2		92.6	93.4	
TOWRE Sight Word Efficiency	83.8	84.7		83.6	84.3	
AIMSWeb (Raw score)	77.3	76.6		76.7	76.9	
WRM Passage Comprehension	91.7	93.6		91.4	94.1	*
GRADE	90.3	91.9		89.9	91.8	*
Woodcock Johnson Spelling	90.8	90.6		89.8	90.1	
Woodcock Johnson Calculation	94.3	94.4		94.3	94.7	
Sample Size	227	181		165	115	

Note: Weights used to account for differential randomization probabilities and nonresponse.

Note: All test scores are shown as standard scores, unless otherwise indicated. All standard scores have mean 100 and standard deviation 15, except for CTOPP and Clinical Evaluation of Language Fundamentals-IV, which have mean 10 and standard deviation 3.

* Difference between treatment and control groups is statistically significant at the 0.05 level.

a Values suppressed to protect student confidentiality.

APPENDIX H

IMPACT ESTIMATE STANDARD ERRORS, TEST STATISTICS, AND P-VALUES

Standard Errors of Impact Estimates for 3rd and 5th Grade Cohorts

Grade 3 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
Word Attack	1.04	1.15	1.90	1.91	1.56	2.33
TOWRE PDE	0.98	1.15	1.78	0.85	1.33	2.82
Word Identification	0.59	0.67	0.98	1.00	1.00	1.36
TOWRE SWE	0.54	0.68	0.87	0.90	1.01	1.67
AIMSweb	2.79	3.54	2.84	2.63	6.74	6.66
Passage Comprehension	0.85	1.06	1.07	1.41	2.40	1.47
GRADE	1.78	2.11	3.45	3.00	5.20	2.78

Grade 5 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
Word Attack	0.67	0.88	0.91	1.96	1.78	0.98
TOWRE PDE	0.90	1.05	1.42	2.00	1.65	1.62
Word Identification	0.68	0.87	0.65	1.23	1.68	1.54
TOWRE SWE	0.86	0.95	1.39	1.15	1.11	1.92
AIMSweb	1.84	1.97	3.86	2.13	3.96	3.53
Passage Comprehension	0.78	0.91	1.34	1.19	1.39	1.78
GRADE	0.96	1.26	1.22	2.23	2.71	1.43

Table H.2

Standard Errors of Impact Estimates for 3rd and 5th Grade Cohorts with Low Baseline Word Attack Scores

Grade 3 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
Word Attack	1.71	2.02	2.98	3.34	3.47	3.72
TOWRE PDE	1.76	2.02	3.25	3.33	3.39	3.73
Word Identification	0.96	1.13	1.61	1.79	1.93	2.05
TOWRE SWE	1.02	1.20	1.76	1.92	2.13	2.08
AIMSweb	3.41	4.02	5.69	6.28	6.69	7.63
Passage Comprehension	1.73	2.02	3.10	3.30	3.25	3.90
GRADE	2.51	2.95	4.19	4.63	5.10	5.44

Grade 5 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
Word Attack	1.37	1.56	2.62	2.39	2.60	2.73
TOWRE PDE	1.36	1.54	2.61	2.36	2.67	2.73
Word Identification	1.07	1.22	2.01	1.86	2.09	2.20
TOWRE SWE	1.00	1.14	1.87	1.70	1.90	2.06
AIMSweb	2.91	3.32	5.44	4.99	5.43	6.05
Passage Comprehension	1.49	1.68	2.84	2.53	2.91	2.96
GRADE	2.06	2.35	3.94	3.56	4.06	4.11

Table H.3

Standard Errors of Impact Estimates for 3rd and 5th Grade Cohorts with High Baseline Word Attack Scores

Grade 3 Cohort	All Interventions	Word-level Interventions	Failure Free Reading	Spell Read	Wilson Reading	Corrective Reading
	Standard Error	Standard Error	Standard Error	Standard Error	Standard Error	Standard Error
Word Attack	2.19	2.55	4.06	3.84	4.45	4.42
TOWRE PDE	0.91	1.07	1.80	1.58	1.81	1.76
Word Identification	1.14	1.33	2.17	1.96	2.30	2.25
TOWRE SWE	0.93	1.07	1.81	1.62	1.83	1.74
AIMSweb	3.11	3.61	6.03	5.31	6.34	6.03
Passage Comprehension	1.20	1.37	2.29	2.13	2.27	2.28
GRADE	2.38	2.76	4.68	3.93	4.82	4.59

Grade 5 Cohort	All Interventions	Word-level Interventions	Failure Free Reading	Spell Read	Wilson Reading	Corrective Reading
	Standard Error	Standard Error	Standard Error	Standard Error	Standard Error	Standard Error
Word Attack	1.14	1.30	2.22	2.29	2.31	1.98
TOWRE PDE	1.08	1.22	2.15	2.20	2.19	1.88
Word Identification	1.11	1.25	2.20	2.20	2.23	1.87
TOWRE SWE	0.95	1.08	1.91	1.93	1.97	1.67
AIMSweb	2.67	3.04	5.35	5.51	5.58	4.66
Passage Comprehension	1.21	1.37	2.39	2.38	2.62	2.03
GRADE	1.55	1.78	2.91	3.22	3.25	2.78

Table H.4

Standard Errors of Impact Estimates for 3rd and 5th Grade Cohorts with Low Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions	Word-level Interventions	Failure Free Reading	Spell Read	Wilson Reading	Corrective Reading
	Standard Error	Standard Error	Standard Error	Standard Error	Standard Error	Standard Error
Word Attack	1.86	2.09	3.41	3.43	3.49	3.38
TOWRE PDE	1.53	1.64	3.01	2.68	2.72	2.45
Word Identification	1.09	1.19	1.94	1.86	1.83	1.88
TOWRE SWE	1.19	1.30	2.11	2.02	1.96	2.09
AIMSweb	4.68	5.22	8.54	8.63	8.95	8.55
Passage Comprehension	1.46	1.60	2.53	2.40	2.49	2.49
GRADE	2.97	3.30	5.24	5.27	5.37	5.32

Grade 5 Cohort	All Interventions	Word-level Interventions	Failure Free Reading	Spell Read	Wilson Reading	Corrective Reading
	Standard Error	Standard Error	Standard Error	Standard Error	Standard Error	Standard Error
Word Attack	1.65	1.90	2.42	2.27	2.89	3.04
TOWRE PDE	1.64	1.90	2.53	2.35	3.13	3.11
Word Identification	1.27	1.46	1.99	1.80	2.34	2.34
TOWRE SWE	1.27	1.46	1.93	1.75	2.26	2.40
AIMSweb	5.26	5.96	7.95	7.12	9.80	9.45
Passage Comprehension	1.47	1.68	2.27	2.05	2.53	2.75
GRADE	2.15	2.48	3.19	2.86	3.72	3.95

Table H.5

Standard Errors of Impact Estimates for 3rd and 5th Grade Cohorts with High Screening Peabody Picture Vocabulary Test

Grade 3 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
Word Attack	1.42	1.67	2.50	2.44	3.04	3.01
TOWRE PDE	1.12	1.31	2.01	1.93	2.36	2.41
Word Identification	0.93	1.09	1.69	1.70	2.03	1.90
TOWRE SWE	1.26	1.48	2.29	2.23	2.65	2.72
AIMSweb	3.04	3.55	5.64	5.44	6.39	6.49
Passage Comprehension	1.54	1.79	2.82	2.86	3.11	3.31
GRADE	3.15	3.67	5.65	5.48	6.32	6.98

Grade 5 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
Word Attack	1.18	1.38	2.27	2.41	2.57	2.17
TOWRE PDE	1.14	1.32	2.22	2.34	2.46	2.06
Word Identification	0.87	1.03	1.55	1.84	1.89	1.60
TOWRE SWE	1.02	1.20	1.97	2.13	2.25	1.83
AIMSweb	2.36	2.80	4.36	5.06	5.26	4.22
Passage Comprehension	1.46	1.68	2.83	2.93	3.25	2.59
GRADE	2.02	2.34	3.87	4.18	4.45	3.60

Table H.6

Standard Errors of Impact Estimates for 3rd and 5th Grade Cohorts with Low Baseline Word Attack and Low Screening Peabody
Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
Word Attack	2.06	2.33	3.62	3.64	4.71	3.95
TOWRE PDE	2.54	2.82	4.86	4.34	5.56	4.56
Word Identification	1.31	1.45	2.31	2.30	2.87	2.51
TOWRE SWE	1.61	1.82	2.84	2.88	3.64	3.16
AIMSweb	4.64	5.27	8.05	8.22	10.58	8.90
Passage Comprehension	1.91	2.15	3.39	3.34	4.29	3.70
GRADE	4.21	5.01	6.80	7.28	13.48	7.92

Grade 5 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
Word Attack	2.18	2.60	3.28	3.10	3.79	4.05
TOWRE PDE	2.11	2.53	3.31	3.11	3.92	4.14
Word Identification	1.44	1.73	2.19	2.00	2.52	2.82
TOWRE SWE	1.70	2.04	2.50	2.45	3.04	3.27
AIMSweb	6.58	7.58	10.78	9.70	11.68	12.40
Passage Comprehension	2.15	2.55	3.29	3.00	3.80	4.10
GRADE	3.18	3.81	4.82	4.57	5.75	6.04

Table H.7

Standard Errors of Impact Estimates for 3rd and 5th Grade Cohorts With Low Baseline Word Attack and High Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
Word Attack	2.34	2.76	3.94	4.28	4.59	5.10
TOWRE PDE	1.81	2.15	3.04	3.28	3.67	4.13
Word Identification	1.42	1.68	2.38	2.67	2.87	3.00
TOWRE SWE	1.99	2.33	3.51	3.72	4.00	4.34
AIMSweb	5.02	5.88	8.97	9.69	9.86	10.51
Passage Comprehension	2.56	3.02	4.29	4.61	4.96	5.91
GRADE	3.47	4.07	5.93	6.76	6.69	7.19

Grade 5 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
Word Attack	1.75	2.00	3.22	3.23	3.40	3.50
TOWRE PDE	1.59	1.81	2.91	2.78	3.22	3.05
Word Identification	1.44	1.67	2.44	2.48	2.96	2.90
TOWRE SWE	1.42	1.62	2.56	2.40	2.91	2.75
AIMSweb	4.44	5.10	8.20	8.33	8.96	8.67
Passage Comprehension	2.35	2.67	4.34	4.24	4.85	4.31
GRADE	3.39	3.87	6.36	6.19	7.20	6.35

Table H.8

Standard Errors of Impact Estimates for 3rd and 5th Grade Cohorts with High Baseline Word Attack and High Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
Word Attack	1.74	2.07	2.98	2.53	4.34	3.02
TOWRE PDE	1.41	1.66	2.42	2.17	3.24	2.56
Word Identification	1.51	1.81	2.55	2.34	3.79	2.72
TOWRE SWE	1.45	1.73	2.47	2.26	3.51	2.60
AIMSweb	4.35	5.11	7.42	6.72	10.08	8.12
Passage Comprehension	1.64	1.91	2.87	2.61	3.70	2.94
GRADE	3.60	4.29	6.13	5.43	8.43	6.75

Grade 5 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
Word Attack	1.54	1.87	2.80	3.39	3.64	2.46
TOWRE PDE	1.50	1.81	2.80	3.40	3.51	2.41
Word Identification	1.32	1.62	2.37	3.14	2.98	2.04
TOWRE SWE	1.33	1.61	2.49	3.00	3.10	2.11
AIMSweb	3.69	4.46	6.94	8.06	8.74	6.01
Passage Comprehension	1.80	2.16	3.42	3.92	4.30	2.85
GRADE	2.08	2.58	3.64	4.88	5.08	3.25

Table H.9

Standard Errors of Impact Estimates for 3rd and 5th Grade Cohorts Eligible for Free or Reduced Price School Lunch

Grade 3 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
Word Attack	1.89	2.04	2.84	2.92	2.77	3.21
TOWRE PDE	1.31	1.39	2.01	1.89	1.84	2.12
Word Identification	1.53	1.70	2.46	2.65	2.39	2.91
TOWRE SWE	1.43	1.54	2.25	2.23	2.15	2.50
AIMSweb	4.52	4.96	7.01	7.50	7.00	8.30
Passage Comprehension	1.55	1.66	2.43	2.37	2.27	2.66
GRADE	3.75	4.17	5.97	6.39	5.94	7.20

Grade 5 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
Word Attack	1.79	2.04	2.74	2.96	3.15	3.09
TOWRE PDE	1.58	1.80	2.43	2.62	2.78	2.78
Word Identification	1.25	1.41	1.85	2.02	2.18	2.08
TOWRE SWE	1.16	1.31	1.65	1.83	1.90	1.88
AIMSweb	3.57	4.01	5.48	5.68	6.08	6.02
Passage Comprehension	1.61	1.80	2.42	2.50	2.76	2.71
GRADE	2.00	2.24	2.86	3.09	3.25	3.32

Table H.10

Standard Errors of Impact Estimates for 3rd and 5th Grade Cohorts Not Eligible for Free or Reduced Price School Lunch

Grade 3 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
Word Attack	1.96	2.44	2.96	2.97	6.01	3.18
TOWRE PDE	1.44	1.81	2.09	2.09	4.45	2.28
Word Identification	1.25	1.58	1.62	1.74	3.99	1.87
TOWRE SWE	1.33	1.68	1.72	1.76	4.34	1.90
AIMSweb	4.16	5.05	7.01	6.89	11.40	7.55
Passage Comprehension	1.74	2.16	2.61	2.63	5.12	2.96
GRADE	3.13	3.86	4.90	4.97	8.69	5.74

Grade 5 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
Word Attack	1.13	1.32	2.30	2.14	2.48	2.27
TOWRE PDE	1.14	1.31	2.35	2.13	2.52	2.20
Word Identification	1.09	1.25	2.30	2.04	2.34	2.12
TOWRE SWE	1.07	1.22	2.26	1.98	2.35	2.01
AIMSweb	2.75	3.15	5.85	5.21	6.04	5.21
Passage Comprehension	1.56	1.77	3.28	2.88	3.51	2.86
GRADE	2.09	2.38	4.39	3.94	4.62	3.96

Table H.11

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts

	All Interventions		Word-level Interventions		Failure Free Reading		Spell Read		Wilson Reading		Corrective Reading	
	ABCD		BCD		A		B		C		D	
Grade 3 Cohort	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	26.53	0.00	22.77	0.00	6.77	0.01	8.11	0.00	13.59	0.00	5.04	0.02
TOWRE PDE	16.38	0.00	17.76	0.00	0.51	0.50	33.24	0.00	9.71	0.00	3.82	0.05
Word Identification	15.28	0.00	13.96	0.00	3.44	0.06	0.48	0.50	17.03	0.00	3.71	0.05
TOWRE SWE	10.05	0.00	5.64	0.02	5.25	0.02	0.90	0.50	6.50	0.01	0.67	0.50
AIMSweb	3.65	0.05	1.60	0.20	7.75	0.01	5.15	0.02	0.28	0.50	0.34	0.50
Passage Comprehension	6.01	0.01	1.53	0.21	17.03	0.00	0.01	0.50	2.13	0.14	0.05	0.50
GRADE	0.30	0.50	0.03	0.50	0.66	0.50	0.48	0.50	0.00	0.50	0.15	0.50

	All Interventions		Word-level Interventions		Failure Free Reading		Spell Read		Wilson Reading		Corrective Reading	
	ABCD		BCD		A		B		C		D	
Grade 5 Cohort	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	15.90	0.00	18.89	0.00	0.75	0.50	3.17	0.07	19.32	0.00	0.03	0.50
TOWRE PDE	3.66	0.05	5.14	0.02	0.03	0.50	2.56	0.11	2.40	0.12	0.72	0.50
Word Identification	0.74	0.50	0.47	0.50	0.75	0.50	0.01	0.50	0.00	0.50	1.50	0.22
TOWRE SWE	2.69	0.10	2.16	0.14	1.09	0.30	8.87	0.00	1.05	0.31	0.04	0.50
AIMSweb	4.51	0.03	3.85	0.05	1.11	0.29	2.43	0.11	0.57	0.50	2.26	0.13
Passage Comprehension	2.13	0.14	0.59	0.50	3.43	0.06	0.54	0.50	0.41	0.50	1.38	0.24
GRADE	0.52	0.50	0.94	0.50	0.54	0.50	0.23	0.50	3.02	0.08	0.00	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
     "p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.12

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts with Low Baseline Word Attack Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	7.22	0.01	6.23	0.01	1.65	0.20	4.99	0.02	0.66	0.50	3.11	0.07
TOWRE PDE	6.05	0.01	5.81	0.02	0.32	0.50	5.93	0.01	0.00	0.50	2.96	0.08
Word Identification	7.44	0.01	5.07	0.02	1.62	0.20	0.06	0.50	21.88	0.00	0.65	0.50
TOWRE SWE	2.01	0.15	0.46	0.50	3.42	0.06	0.41	0.50	1.55	0.21	1.27	0.26
AIMSweb	0.59	0.50	0.00	0.50	6.62	0.01	4.24	0.04	0.19	0.50	0.20	0.50
Passage Comprehension	1.47	0.22	0.78	0.50	1.52	0.22	0.52	0.50	1.47	0.22	1.64	0.20
GRADE	0.31	0.50	0.35	0.50	0.02	0.50	1.10	0.29	2.66	0.10	1.18	0.28

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	21.05	0.00	27.20	0.00	0.35	0.50	18.04	0.00	6.68	0.01	5.10	0.02
TOWRE PDE	2.27	0.13	4.55	0.03	0.82	0.50	2.25	0.13	1.48	0.22	2.30	0.13
Word Identification	0.00	0.50	0.02	0.50	0.21	0.50	0.03	0.50	0.36	0.50	0.07	0.50
TOWRE SWE	4.75	0.03	3.21	0.07	4.38	0.03	5.23	0.02	0.64	0.50	0.42	0.50
AIMSweb	4.28	0.04	3.77	0.05	1.32	0.25	2.61	0.10	4.21	0.04	0.22	0.50
Passage Comprehension	0.29	0.50	0.85	0.50	0.78	0.50	0.04	0.50	2.09	0.14	0.02	0.50
GRADE	4.75	0.03	5.95	0.01	0.08	0.50	0.14	0.50	11.03	0.00	2.45	0.11

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
   "p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.13

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts with High Baseline Word Attack Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	13.86	0.00	10.77	0.00	3.99	0.04	1.39	0.24	4.23	0.04	5.15	0.02
TOWRE PDE	36.89	0.00	54.85	0.00	0.00	0.50	39.69	0.00	51.66	0.00	5.34	0.02
Word Identification	15.22	0.00	9.90	0.00	5.27	0.02	0.35	0.50	6.66	0.01	3.64	0.05
TOWRE SWE	32.37	0.00	17.98	0.00	22.90	0.00	1.05	0.31	6.49	0.01	9.74	0.00
AIMSweb	4.23	0.04	4.82	0.03	0.02	0.50	2.35	0.12	5.02	0.02	0.49	0.50
Passage Comprehension	5.17	0.02	0.56	0.50	11.10	0.00	0.94	0.50	0.26	0.50	0.80	0.50
GRADE	0.08	0.50	1.40	0.23	4.74	0.03	3.21	0.07	14.05	0.00	0.16	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	1.03	0.31	1.50	0.22	0.01	0.50	0.49	0.50	22.89	0.00	0.48	0.50
TOWRE PDE	2.49	0.11	1.49	0.22	0.89	0.50	0.95	0.50	0.39	0.50	0.04	0.50
Word Identification	4.75	0.03	2.81	0.09	3.01	0.08	1.68	0.19	0.18	0.50	2.56	0.11
TOWRE SWE	1.82	0.17	1.80	0.18	0.24	0.50	15.44	0.00	0.18	0.50	0.04	0.50
AIMSweb	0.46	0.50	0.26	0.50	0.24	0.50	1.32	0.25	0.01	0.50	2.64	0.10
Passage Comprehension	5.60	0.02	3.41	0.06	2.29	0.13	1.91	0.16	0.00	0.50	4.11	0.04
GRADE	2.97	0.08	0.71	0.50	6.00	0.01	0.01	0.50	0.14	0.50	4.20	0.04

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
   "p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.14

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts with Low Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	10.59	0.00	5.11	0.02	6.31	0.01	0.06	0.50	7.21	0.01	4.85	0.03
TOWRE PDE	7.51	0.01	8.14	0.00	2.80	0.09	0.61	0.50	5.95	0.01	8.79	0.00
Word Identification	2.47	0.11	2.48	0.11	0.89	0.50	0.27	0.50	5.62	0.02	1.71	0.19
TOWRE SWE	1.53	0.21	1.49	0.22	0.26	0.50	0.31	0.50	3.16	0.07	0.02	0.50
AIMSweb	0.25	0.50	1.53	0.21	2.38	0.12	0.11	0.50	3.06	0.08	0.01	0.50
Passage Comprehension	0.44	0.50	0.02	0.50	1.26	0.26	0.06	0.50	0.60	0.50	0.26	0.50
GRADE	0.39	0.50	0.98	0.50	0.17	0.50	2.03	0.15	0.75	0.50	0.94	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	7.47	0.01	13.81	0.00	2.65	0.10	0.36	0.50	18.04	0.00	2.69	0.10
TOWRE PDE	8.95	0.00	7.59	0.01	1.42	0.23	0.23	0.50	2.26	0.13	3.32	0.06
Word Identification	0.04	0.50	0.12	0.50	0.12	0.50	0.97	0.50	0.79	0.50	0.34	0.50
TOWRE SWE	8.45	0.00	9.16	0.00	0.32	0.50	7.25	0.01	1.87	0.17	4.98	0.02
AIMSweb	2.64	0.10	1.26	0.26	3.20	0.07	1.95	0.16	0.00	0.50	0.91	0.50
Passage Comprehension	3.10	0.07	4.49	0.03	0.02	0.50	1.48	0.22	8.44	0.00	0.91	0.50
GRADE	5.48	0.02	3.97	0.04	4.41	0.03	0.04	0.50	5.24	0.02	4.37	0.03

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
    "p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.15

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts with High Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	11.92	0.00	11.32	0.00	1.59	0.21	18.65	0.00	2.33	0.12	5.05	0.02
TOWRE PDE	9.81	0.00	8.24	0.00	1.34	0.25	47.17	0.00	2.25	0.13	0.00	0.50
Word Identification	12.42	0.00	8.88	0.00	5.35	0.02	1.31	0.25	15.58	0.00	0.88	0.50
TOWRE SWE	4.26	0.04	1.90	0.17	6.12	0.01	1.17	0.28	1.42	0.23	0.28	0.50
AIMSweb	14.25	0.00	10.87	0.00	7.22	0.01	4.95	0.02	16.65	0.00	0.44	0.50
Passage Comprehension	5.51	0.02	1.77	0.18	17.11	0.00	0.07	0.50	3.25	0.07	0.17	0.50
GRADE	1.86	0.17	0.56	0.50	2.58	0.10	0.90	0.50	0.15	0.50	1.05	0.31

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	4.82	0.03	6.07	0.01	0.19	0.50	7.91	0.01	3.08	0.08	0.52	0.50
TOWRE PDE	5.81	0.02	12.45	0.00	0.40	0.50	21.60	0.00	2.24	0.13	0.71	0.50
Word Identification	0.78	0.50	0.54	0.50	0.46	0.50	1.97	0.16	1.41	0.23	1.73	0.19
TOWRE SWE	1.10	0.29	0.57	0.50	0.85	0.50	4.35	0.03	0.80	0.50	1.67	0.19
AIMSweb	9.36	0.00	8.13	0.00	1.48	0.22	0.88	0.50	1.92	0.16	11.20	0.00
Passage Comprehension	3.40	0.06	1.83	0.17	1.84	0.17	0.59	0.50	0.71	0.50	0.55	0.50
GRADE	0.03	0.50	0.06	0.50	1.71	0.19	0.08	0.50	1.18	0.28	0.78	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
"p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.16

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts with Low Baseline Word Attack and Low Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	11.31	0.00	11.75	0.00	0.21	0.50	2.36	0.12	3.30	0.07	4.66	0.03
TOWRE PDE	0.98	0.50	0.59	0.50	0.36	0.50	1.29	0.25	2.52	0.11	1.51	0.22
Word Identification	0.94	0.50	0.32	0.50	0.62	0.50	0.77	0.50	0.73	0.50	2.19	0.14
TOWRE SWE	0.30	0.50	0.22	0.50	0.11	0.50	3.53	0.06	1.48	0.22	0.08	0.50
AIMSweb	5.53	0.02	7.81	0.01	1.43	0.23	0.48	0.50	13.06	0.00	0.74	0.50
Passage Comprehension	0.30	0.50	0.00	0.50	1.52	0.22	1.06	0.30	0.30	0.50	0.03	0.50
GRADE	0.13	0.50	0.39	0.50	3.40	0.06	0.77	0.50	1.08	0.30	0.70	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	8.91	0.00	12.23	0.00	2.12	0.14	6.49	0.01	2.19	0.14	7.83	0.01
TOWRE PDE	6.87	0.01	6.40	0.01	0.77	0.50	0.26	0.50	3.18	0.07	2.52	0.11
Word Identification	0.94	0.50	0.32	0.50	0.62	0.50	0.77	0.50	0.73	0.50	2.19	0.14
TOWRE SWE	4.48	0.03	3.50	0.06	2.49	0.11	5.44	0.02	0.81	0.50	1.63	0.20
AIMSweb	3.79	0.05	4.28	0.04	0.49	0.50	2.00	0.15	1.13	0.29	3.67	0.05
Passage Comprehension	9.16	0.00	9.23	0.00	0.00	0.50	17.27	0.00	1.07	0.30	4.89	0.03
GRADE	1.03	0.31	0.32	0.50	3.26	0.07	0.57	0.50	2.14	0.14	0.51	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
  "p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.17

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts With Low Baseline Word Attack and High Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	0.00	0.50	0.01	0.50	0.01	0.50	2.02	0.15	0.61	0.50	6.15	0.01
TOWRE PDE	2.45	0.11	1.70	0.19	0.49	0.50	14.62	0.00	1.33	0.25	0.65	0.50
Word Identification	5.50	0.02	3.50	0.06	3.35	0.06	1.31	0.25	25.55	0.00	0.07	0.50
TOWRE SWE	0.96	0.50	0.48	0.50	2.21	0.13	0.01	0.50	3.38	0.06	0.19	0.50
AIMSweb	5.76	0.02	4.13	0.04	4.10	0.04	2.73	0.09	3.62	0.05	0.14	0.50
Passage Comprehension	3.83	0.05	2.23	0.13	8.51	0.00	0.20	0.50	3.24	0.07	1.34	0.25
GRADE	1.46	0.22	0.12	0.50	5.34	0.02	0.83	0.50	0.79	0.50	0.10	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	14.84	0.00	17.44	0.00	0.55	0.50	6.98	0.01	11.68	0.00	0.39	0.50
TOWRE PDE	12.20	0.00	25.00	0.00	0.80	0.50	18.29	0.00	4.91	0.03	3.95	0.04
Word Identification	0.08	0.50	0.04	0.50	4.19	0.04	2.06	0.15	0.00	0.50	0.43	0.50
TOWRE SWE	6.09	0.01	5.65	0.02	0.87	0.50	9.99	0.00	1.87	0.17	0.26	0.50
AIMSweb	3.03	0.08	1.57	0.21	3.57	0.06	0.08	0.50	5.20	0.02	0.02	0.50
Passage Comprehension	0.18	0.50	0.10	0.50	0.07	0.50	3.44	0.06	0.88	0.50	0.08	0.50
GRADE	5.94	0.01	8.71	0.00	0.09	0.50	0.56	0.50	6.08	0.01	2.63	0.10

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
"p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.18

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts with High Baseline Word Attack and High Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	59.88	0.00	49.59	0.00	3.81	0.05	64.26	0.00	3.86	0.05	10.49	0.00
TOWRE PDE	10.36	0.00	14.56	0.00	0.95	0.50	19.70	0.00	5.61	0.02	1.37	0.24
Word Identification	4.01	0.04	6.21	0.01	0.41	0.50	1.69	0.19	0.24	0.50	7.11	0.01
TOWRE SWE	19.19	0.00	12.51	0.00	8.16	0.00	0.01	0.50	6.71	0.01	8.93	0.00
AIMSweb	3.02	0.08	3.63	0.05	0.00	0.50	1.02	0.31	10.23	0.00	0.00	0.50
Passage Comprehension	7.11	0.01	0.70	0.50	17.88	0.00	0.03	0.50	0.20	0.50	0.70	0.50
GRADE	0.25	0.50	0.21	0.50	0.07	0.50	2.18	0.14	1.12	0.29	0.14	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	1.02	0.31	1.20	0.27	0.00	0.50	1.04	0.31	1.43	0.23	0.74	0.50
TOWRE PDE	4.55	0.03	4.42	0.03	0.20	0.50	1.62	0.20	3.03	0.08	0.78	0.50
Word Identification	2.80	0.09	2.71	0.10	0.10	0.50	4.84	0.03	0.02	0.50	1.22	0.27
TOWRE SWE	2.27	0.13	1.25	0.26	1.27	0.26	0.45	0.50	2.52	0.11	0.63	0.50
AIMSweb	0.70	0.50	0.73	0.50	0.02	0.50	0.08	0.50	0.23	0.50	9.38	0.00
Passage Comprehension	0.46	0.50	0.03	0.50	2.19	0.13	0.01	0.50	0.03	0.50	0.25	0.50
GRADE	6.68	0.01	4.11	0.04	6.46	0.01	1.56	0.21	0.35	0.50	7.13	0.01

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
"p"is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.19

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts Eligible for Free or Reduced Price School Lunch

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	9.30	0.00	4.17	0.04	7.65	0.01	0.49	0.50	10.85	0.00	0.83	0.50
TOWRE PDE	6.07	0.01	7.08	0.01	1.22	0.27	16.30	0.00	7.38	0.01	0.16	0.50
Word Identification	3.60	0.05	3.51	0.06	1.34	0.24	0.02	0.50	7.48	0.01	2.77	0.09
TOWRE SWE	5.05	0.02	2.36	0.12	5.54	0.02	0.87	0.50	4.61	0.03	0.00	0.50
AIMSweb	0.75	0.50	0.21	0.50	1.40	0.23	1.21	0.27	0.00	0.50	0.05	0.50
Passage Comprehension	0.02	0.50	0.01	0.50	0.01	0.50	0.06	0.50	0.03	0.50	0.00	0.50
GRADE	0.16	0.50	0.04	0.50	1.30	0.25	0.00	0.50	0.07	0.50	0.70	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	7.41	0.01	8.30	0.00	0.00	0.50	1.27	0.26	5.36	0.02	2.16	0.14
TOWRE PDE	1.80	0.18	2.07	0.15	0.00	0.50	0.06	0.50	0.02	0.50	5.96	0.01
Word Identification	0.03	0.50	0.01	0.50	0.78	0.50	0.43	0.50	0.08	0.50	0.11	0.50
TOWRE SWE	9.44	0.00	12.54	0.00	0.76	0.50	15.91	0.00	1.82	0.17	10.86	0.00
AIMSweb	0.51	0.50	3.18	0.07	2.76	0.09	1.12	0.29	1.07	0.30	1.30	0.25
Passage Comprehension	1.30	0.25	0.15	0.50	3.38	0.06	1.82	0.17	0.40	0.50	0.04	0.50
GRADE	9.04	0.00	4.88	0.03	12.91	0.00	24.78	0.00	0.03	0.50	0.03	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
    "p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.20

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts Not Eligible for Free or Reduced Price School Lunch

| | All Interventions | | Word-level Interventions | | Failure Free Reading | | Spell Read | | Wilson Reading | | Corrective Reading | |
| | ABCD | | BCD | | A | | B | | C | | D | |
Grade 3 Cohort	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*
Word Attack	6.65	0.01	6.30	0.01	0.33	0.50	7.43	0.01	1.29	0.25	5.99	0.01
TOWRE PDE	68.55	0.00	88.03	0.00	0.07	0.50	18.23	0.00	77.03	0.00	24.12	0.00
Word Identification	7.29	0.01	4.21	0.04	6.06	0.01	0.62	0.50	2.15	0.14	1.56	0.21
TOWRE SWE	3.93	0.04	2.57	0.10	4.57	0.03	0.01	0.50	1.17	0.28	2.66	0.10
AIMSweb	7.76	0.01	5.63	0.02	9.79	0.00	0.15	0.50	5.10	0.02	0.46	0.50
Passage Comprehension	5.91	0.01	2.81	0.09	21.48	0.00	0.07	0.50	2.82	0.09	0.50	0.50
GRADE	0.02	0.50	0.00	0.50	0.05	0.50	0.55	0.50	0.57	0.50	0.37	0.50

| | All Interventions | | Word-level Interventions | | Failure Free Reading | | Spell Read | | Wilson Reading | | Corrective Reading | |
| | ABCD | | BCD | | A | | B | | C | | D | |
Grade 5 Cohort	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*
Word Attack	10.95	0.00	14.84	0.00	0.42	0.50	9.52	0.00	18.02	0.00	0.50	0.50
TOWRE PDE	5.11	0.02	8.28	0.00	0.00	0.50	9.93	0.00	4.09	0.04	0.04	0.50
Word Identification	1.01	0.32	1.15	0.28	0.00	0.50	0.01	0.50	0.03	0.50	2.79	0.09
TOWRE SWE	2.67	0.10	1.49	0.22	2.37	0.12	2.87	0.09	3.85	0.05	0.68	0.50
AIMSweb	3.06	0.08	4.32	0.04	0.00	0.50	5.80	0.02	0.14	0.50	8.18	0.00
Passage Comprehension	1.46	0.22	0.83	0.50	0.76	0.50	0.08	0.50	0.38	0.50	0.44	0.50
GRADE	0.30	0.50	0.39	0.50	0.00	0.50	0.36	0.50	0.45	0.50	1.10	0.29

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
   "*p*" is the *p*-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.21

Tests for Differences in Impacts for 3rd and 5th Grade Cohorts

	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	4.67	0.03	1.31	0.25	7.40	0.01	0.50	0.48	0.75	0.38	4.01	0.05
TOWRE PDE	2.85	0.09	2.54	0.11	0.45	0.50	0.63	0.43	0.55	0.46	1.62	0.20
Word Identification	10.32	0.00	7.93	0.00	4.10	0.04	0.15	0.70	4.42	0.04	4.81	0.03
TOWRE SWE	0.08	0.77	0.03	0.85	0.11	0.74	3.10	0.08	0.94	0.33	0.48	0.49
AIMSweb	7.64	0.01	4.24	0.04	6.24	0.01	7.54	0.01	0.70	0.40	1.49	0.22
Passage Comprehension	7.79	0.01	2.07	0.15	16.15	0.00	0.28	0.59	0.89	0.35	1.09	0.30
GRADE	0.02	0.89	0.12	0.72	1.02	0.31	0.70	0.40	0.63	0.43	0.12	0.73

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
   "p"is the p -value produced by HLM6 software

Table H.22

Tests for Differences in Impacts for 3rd and 5th Grade Cohorts with Low Baseline Word Attack Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	1.03	0.31	0.14	0.50	1.74	0.18	0.11	0.50	0.16	0.50	0.51	0.50
TOWRE PDE	0.03	0.50	0.00	0.50	0.23	0.50	0.00	0.50	7.42	0.01	1.39	0.24
Word Identification	1.82	0.17	1.24	0.27	0.29	0.50	0.09	0.50	0.52	0.50	0.52	0.50
TOWRE SWE	5.51	0.02	4.62	0.03	0.84	0.50	0.00	0.50	0.05	0.50	28.44	0.00
AIMSweb	0.52	0.50	2.85	0.09	1.73	0.19	0.70	0.50	2.60	0.10	2.57	0.10
Passage Comprehension	0.06	0.50	0.12	0.50	1.67	0.19	0.86	0.50	1.30	0.25	0.10	0.50
GRADE	0.50	0.50	2.55	0.11	1.79	0.18	7.53	0.01	19.58	0.00	0.05	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	5.58	0.02	8.25	0.00	0.06	0.50	7.67	0.01	0.23	0.50	3.06	0.08
TOWRE PDE	0.42	0.50	2.21	0.13	2.60	0.10	0.32	0.50	0.49	0.50	2.98	0.08
Word Identification	1.56	0.21	0.71	0.50	3.23	0.07	0.95	0.50	0.48	0.50	1.57	0.21
TOWRE SWE	1.10	0.29	0.57	0.50	1.04	0.31	0.12	0.50	0.11	0.50	0.47	0.50
AIMSweb	1.47	0.22	1.22	0.27	0.45	0.50	2.96	0.08	2.10	0.14	0.75	0.50
Passage Comprehension	3.06	0.08	2.85	0.09	0.36	0.50	1.08	0.30	0.91	0.50	1.09	0.30
GRADE	6.48	0.01	5.89	0.01	1.13	0.29	0.14	0.50	7.91	0.01	4.51	0.03

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
    "p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.23

Tests for Differences in Impacts for 3rd and 5th Grade Cohorts with Low Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	0.03	0.50	0.68	0.50	2.00	0.15	4.35	0.03	0.01	0.50	0.68	0.50
TOWRE PDE	0.25	0.50	0.23	0.50	0.11	0.50	1.89	0.17	0.14	0.50	6.05	0.01
Word Identification	0.33	0.50	0.50	0.50	0.00	0.50	1.67	0.19	0.02	0.50	0.00	0.50
TOWRE SWE	0.65	0.50	0.22	0.50	0.91	0.50	0.07	0.50	0.02	0.50	0.15	0.50
AIMSweb	5.72	0.02	8.10	0.00	0.22	0.50	2.09	0.14	7.43	0.01	0.49	0.50
Passage Comprehension	1.81	0.17	0.85	0.50	1.99	0.16	0.08	0.50	3.68	0.05	0.03	0.50
GRADE	0.72	0.50	0.00	0.50	4.06	0.04	0.04	0.50	2.28	0.13	2.17	0.14

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	0.03	0.50	0.21	0.50	0.39	0.50	2.68	0.10	0.40	0.50	2.79	0.09
TOWRE PDE	0.03	0.50	0.91	0.50	0.89	0.50	5.62	0.02	0.28	0.50	1.86	0.17
Word Identification	0.12	0.50	0.03	0.50	0.35	0.50	13.35	0.00	1.79	0.18	0.01	0.50
TOWRE SWE	1.00	0.32	1.63	0.20	0.28	0.50	0.21	0.50	0.07	0.50	4.98	0.02
AIMSweb	0.04	0.50	0.08	0.50	2.41	0.12	0.38	0.50	0.84	0.50	0.01	0.50
Passage Comprehension	5.01	0.02	5.16	0.02	0.45	0.50	1.01	0.32	4.10	0.04	1.58	0.21
GRADE	2.28	0.13	0.81	0.50	4.89	0.03	0.04	0.50	0.00	0.50	3.18	0.07

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
    "p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.24

Tests for Differences in Impacts for 3rd and 5th Grade Cohorts with Low Baseline Word Attack and Low Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	0.61	0.50	0.00	0.50	2.93	0.08	0.00	0.50	0.71	0.50	0.98	0.50
TOWRE PDE	0.41	0.50	0.64	0.50	0.01	0.50	5.77	0.02	0.00	0.50	1.07	0.30
Word Identification	3.11	0.07	3.18	0.07	0.13	0.50	1.46	0.23	1.96	0.16	0.18	0.50
TOWRE SWE	4.15	0.04	2.48	0.11	3.37	0.06	1.35	0.24	3.60	0.05	0.68	0.50
AIMSweb	19.23	0.00	17.77	0.00	0.55	0.50	3.30	0.07	14.10	0.00	0.00	0.50
Passage Comprehension	2.43	0.12	1.37	0.24	1.29	0.25	0.71	0.50	1.33	0.25	0.00	0.50
GRADE	2.03	0.15	0.10	0.50	12.45	0.00	0.02	0.50	2.40	0.12	0.98	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	0.07	0.50	0.52	0.50	0.56	0.50	0.82	0.50	7.96	0.01	6.95	0.01
TOWRE PDE	0.73	0.50	0.23	0.50	1.06	0.30	1.78	0.18	0.11	0.50	1.84	0.17
Word Identification	3.11	0.07	3.18	0.07	0.13	0.50	1.46	0.23	1.96	0.16	0.18	0.50
TOWRE SWE	1.07	0.30	0.95	0.50	0.22	0.50	0.01	0.50	0.14	0.50	1.82	0.17
AIMSweb	1.89	0.17	2.36	0.12	0.04	0.50	1.54	0.21	0.33	0.50	1.93	0.16
Passage Comprehension	12.88	0.00	12.19	0.00	1.30	0.25	17.05	0.00	0.58	0.50	7.73	0.01
GRADE	0.54	0.50	0.01	0.50	6.95	0.01	0.65	0.50	0.01	0.50	0.55	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
    "p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.25

Tests for Differences in Impacts for 3rd and 5th Grade Cohorts With Low Baseline Word Attack and High Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	4.10	0.04	3.01	0.08	1.37	0.24	0.10	0.50	0.09	0.50	10.78	0.00
TOWRE PDE	2.83	0.09	6.08	0.01	0.79	0.50	0.32	0.50	0.80	0.50	9.68	0.00
Word Identification	0.00	0.50	0.02	0.50	0.09	0.50	2.05	0.15	1.55	0.21	0.96	0.50
TOWRE SWE	0.02	0.50	0.01	0.50	0.01	0.50	0.10	0.50	1.05	0.31	0.77	0.50
AIMSweb	1.99	0.15	2.33	0.12	0.00	0.50	1.29	0.25	2.07	0.15	0.16	0.50
Passage Comprehension	0.72	0.50	1.11	0.29	0.20	0.50	0.33	0.50	2.87	0.09	0.20	0.50
GRADE	1.33	0.25	0.05	0.50	8.86	0.00	1.30	0.25	1.25	0.26	0.19	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	3.45	0.06	4.18	0.04	0.02	0.50	3.03	0.08	1.42	0.23	0.26	0.50
TOWRE PDE	3.44	0.06	10.01	0.00	1.87	0.17	7.17	0.01	2.25	0.13	0.62	0.50
Word Identification	0.12	0.50	0.81	0.50	4.55	0.03	8.59	0.00	0.01	0.50	0.03	0.50
TOWRE SWE	0.85	0.50	0.95	0.50	0.08	0.50	0.78	0.50	0.85	0.50	0.02	0.50
AIMSweb	0.72	0.50	0.27	0.50	1.14	0.28	0.03	0.50	4.16	0.04	0.34	0.50
Passage Comprehension	0.07	0.50	0.03	0.50	0.04	0.50	1.19	0.27	0.76	0.50	0.08	0.50
GRADE	7.54	0.01	10.67	0.00	0.01	0.50	1.48	0.22	5.20	0.02	3.44	0.06

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
    "p "is the p -value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.26

Tests for Differences in Impacts for 3rd and 5th Grade Cohorts with High Baseline Word Attack and High Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	1.88	0.17	2.98	0.08	0.19	0.50	12.13	0.00	0.04	0.50	0.09	0.50
TOWRE PDE	0.31	0.50	0.14	0.50	0.34	0.50	0.09	0.50	0.00	0.50	0.48	0.50
Word Identification	0.31	0.50	0.00	0.50	1.76	0.18	0.18	0.50	1.72	0.19	0.61	0.50
TOWRE SWE	3.43	0.06	3.24	0.07	0.20	0.50	0.20	0.50	0.23	0.50	19.42	0.00
AIMSweb	0.02	0.50	1.07	0.30	3.10	0.07	0.08	0.50	3.33	0.06	0.08	0.50
Passage Comprehension	0.46	0.50	0.02	0.50	3.00	0.08	0.54	0.50	0.46	0.50	0.05	0.50
GRADE	0.10	0.50	0.01	0.50	0.65	0.50	1.97	0.16	2.68	0.10	0.29	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	2.64	0.10	3.83	0.05	0.10	0.50	0.29	0.50	1.27	0.26	2.78	0.09
TOWRE PDE	0.01	0.50	0.07	0.50	0.40	0.50	0.03	0.50	0.01	0.50	0.62	0.50
Word Identification	1.53	0.21	1.64	0.20	0.03	0.50	1.61	0.20	0.16	0.50	0.67	0.50
TOWRE SWE	0.32	0.50	0.52	0.50	0.13	0.50	3.18	0.07	0.58	0.50	2.40	0.12
AIMSweb	0.13	0.50	0.00	0.50	0.43	0.50	1.38	0.24	1.27	0.26	4.13	0.04
Passage Comprehension	0.33	0.50	0.22	0.50	0.13	0.50	0.00	0.50	0.84	0.50	0.00	0.50
GRADE	11.15	0.00	9.30	0.00	2.92	0.08	1.98	0.16	3.67	0.05	8.65	0.00

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
 "p"is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.27

Tests for Differences in Impacts for 3rd and 5th Grade Cohorts Eligible for Free or Reduced Price School Lunch

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	0.79	0.50	2.23	0.13	2.75	0.09	1.97	0.16	0.44	0.50	0.93	0.50
TOWRE PDE	13.02	0.00	25.75	0.00	0.42	0.50	0.01	0.50	35.69	0.00	16.66	0.00
Word Identification	0.20	0.50	0.36	0.50	0.02	0.50	0.14	0.50	0.30	0.50	1.03	0.31
TOWRE SWE	0.00	0.50	0.19	0.50	1.46	0.23	0.68	0.50	0.07	0.50	1.87	0.17
AIMSweb	2.47	0.11	2.58	0.10	0.01	0.50	0.06	0.50	2.93	0.08	0.37	0.50
Passage Comprehension	5.79	0.02	3.01	0.08	6.78	0.01	0.00	0.50	3.07	0.08	0.26	0.50
GRADE	0.07	0.50	0.04	0.50	0.80	0.50	0.30	0.50	0.82	0.50	1.15	0.28

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
Word Attack	0.02	0.50	0.12	0.50	0.20	0.50	0.15	0.50	1.05	0.31	1.94	0.16
TOWRE PDE	0.62	0.50	1.02	0.31	0.00	0.50	2.57	0.10	2.68	0.10	3.17	0.07
Word Identification	0.49	0.50	0.98	0.50	0.50	0.50	0.05	0.50	0.11	0.50	1.19	0.28
TOWRE SWE	0.71	0.50	1.14	0.28	0.21	0.50	0.06	0.50	0.26	0.50	4.73	0.03
AIMSweb	3.23	0.07	6.68	0.01	2.20	0.13	7.26	0.01	0.57	0.50	4.79	0.03
Passage Comprehension	0.04	0.50	0.19	0.50	0.20	0.50	0.19	0.50	0.94	0.50	0.11	0.50
GRADE	3.12	0.07	2.25	0.13	1.12	0.29	6.00	0.01	0.20	0.50	0.61	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
"p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.28

Standard Errors of Impact Estimates for 3rd and 5th Grade Cohorts

Grade 3 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
PSSA Reading	16.89	20.70	22.99	21.28	47.35	28.18
PSSA Math	18.15	15.98	43.64	30.05	21.54	19.08

Grade 5 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
PSSA Reading	11.77	14.65	13.53	16.92	29.04	30.23
PSSA Math	8.56	10.28	10.87	18.72	18.07	12.43

Table H.29

Standard Errors of Impact Estimates for 3rd and 5th Grade Cohorts with Low Baseline Word Attack Scores

Grade 3 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
PSSA Reading	27.93	33.30	44.69	46.15	56.25	70.78
PSSA Math	32.04	38.18	49.73	51.14	69.13	77.86

Grade 5 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
PSSA Reading	21.28	24.15	40.32	35.35	41.50	40.20
PSSA Math	23.10	26.26	43.54	40.23	45.20	43.84

Table H.30

Standard Errors of Impact Estimates for 3rd and 5th Grade Cohorts with High Baseline Word Attack Scores

Grade 3 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
PSSA Reading	22.87	26.31	44.17	37.61	47.74	42.27
PSSA Math	21.80	25.01	44.47	38.03	43.20	40.73
Grade 5 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
PSSA Reading	27.74	31.19	55.86	51.66	59.66	49.10
PSSA Math	24.25	27.63	47.69	47.30	51.19	43.47

Table H.31

Standard Errors of Impact Estimates for 3rd and 5th Grade Cohorts with Low Screening Peabody Picture Vocabulary Test

Grade 3 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
PSSA Reading	34.14	36.41	55.12	49.44	55.10	57.75
PSSA Math	47.59	52.91	82.63	80.88	86.50	89.34

Grade 5 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
PSSA Reading	29.66	34.13	43.38	38.28	52.33	55.03
PSSA Math	30.01	34.45	44.12	38.46	52.67	54.88

Table H.32

Standard Errors of Impact Estimates for 3rd and 5th Grade Cohorts with High Screening Peabody Picture Vocabulary Test

Grade 3 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
PSSA Reading	17.70	20.75	31.97	32.48	40.87	34.77
PSSA Math	19.45	22.93	33.76	34.94	44.90	38.58

Grade 5 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
PSSA Reading	22.29	26.16	41.87	46.02	49.82	40.05
PSSA Math	18.74	22.26	34.09	39.56	41.71	34.03

Table H.33

Standard Errors of Impact Estimates for 3rd and 5th Grade Cohorts with Low Baseline Word Attack and Low Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
PSSA Reading	40.02	44.24	64.46	61.58	87.25	78.74
PSSA Math	60.15	68.31	100.33	99.42	136.83	123.52

Grade 5 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
PSSA Reading	42.17	49.91	64.21	58.25	73.79	77.05
PSSA Math	43.09	51.42	64.55	59.10	75.71	78.11

Table H.34

Standard Errors of Impact Estimates for 3rd and 5th Grade Cohorts With Low Baseline Word Attack and High Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
PSSA Reading	31.56	37.36	51.84	56.16	66.80	70.60
PSSA Math	30.90	36.96	49.26	56.26	67.54	67.08

Grade 5 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
PSSA Reading	29.54	33.81	56.98	58.27	60.90	57.45
PSSA Math	28.99	33.08	55.19	57.85	59.03	55.55

Table H.35

Standard Errors of Impact Estimates for 3rd and 5th Grade Cohorts with High Baseline Word Attack and High Screening
Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
PSSA Reading	27.17	31.61	47.55	43.40	64.60	47.24
PSSA Math	32.44	38.81	54.94	50.67	82.65	55.05

Grade 5 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
PSSA Reading	34.84	41.71	66.55	74.89	82.27	57.85
PSSA Math	29.84	36.68	52.10	66.04	72.66	47.27

Table H.36

Standard Errors of Impact Estimates for 3rd and 5th Grade Cohorts Eligible for Free or Reduced Price School Lunch

Grade 3 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
PSSA Reading	27.42	30.30	43.69	44.65	42.23	48.40
PSSA Math	30.48	33.72	45.97	49.04	46.77	53.88

Grade 5 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
PSSA Reading	29.80	32.95	39.73	44.25	46.23	43.35
PSSA Math	29.58	32.96	40.81	46.30	47.30	44.65

Table H.37

Standard Errors of Impact Estimates for 3rd and 5th Grade Cohorts Not Eligible for Free or Reduced Price School Lunch

Grade 3 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
PSSA Reading	30.23	37.56	43.91	45.50	88.33	53.78
PSSA Math	35.92	45.21	47.41	48.14	111.83	55.52

Grade 5 Cohort	All Interventions Standard Error	Word-level Interventions Standard Error	Failure Free Reading Standard Error	Spell Read Standard Error	Wilson Reading Standard Error	Corrective Reading Standard Error
PSSA Reading	23.24	27.36	45.47	43.52	54.74	44.68
PSSA Math	23.25	27.07	46.94	43.62	51.40	45.97

Table H.38

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	0.85	0.50	0.03	0.50	4.94	0.02	3.51	0.06	1.23	0.27	0.71	0.50
PSSA Math	1.24	0.26	0.79	0.50	0.78	0.50	0.27	0.50	6.91	0.01	0.01	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	5.39	0.02	2.99	0.08	6.09	0.01	3.15	0.07	0.67	0.50	0.54	0.50
PSSA Math	11.34	0.00	10.91	0.00	1.51	0.22	1.16	0.28	9.75	0.00	4.17	0.04

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
"p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.39

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts with Low Baseline Word Attack Scores

	All Interventions		Word-level Interventions		Failure Free Reading		Spell Read		Wilson Reading		Corrective Reading	
	ABCD		BCD		A		B		C		D	
Grade 3 Cohort	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	0.01	0.50	1.22	0.27	9.52	0.00	0.16	0.50	0.94	0.50	1.03	0.31
PSSA Math	4.64	0.03	5.76	0.02	0.36	0.50	0.56	0.50	8.50	0.00	4.35	0.03

	All Interventions		Word-level Interventions		Failure Free Reading		Spell Read		Wilson Reading		Corrective Reading	
	ABCD		BCD		A		B		C		D	
Grade 5 Cohort	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	2.92	0.08	1.39	0.24	2.85	0.09	1.65	0.20	2.01	0.15	0.10	0.50
PSSA Math	0.68	0.50	1.41	0.23	0.29	0.50	0.29	0.50	1.61	0.20	0.18	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
   "p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.40

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts with High Baseline Word Attack Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	2.35	0.12	2.64	0.10	0.17	0.50	4.61	0.03	0.59	0.50	2.67	0.10
PSSA Math	0.07	0.50	1.61	0.20	0.80	0.50	0.12	0.50	1.41	0.23	22.66	0.00

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	4.27	0.04	3.20	0.07	0.81	0.50	0.75	0.50	0.23	0.50	1.92	0.16
PSSA Math	4.56	0.03	3.95	0.04	0.68	0.50	0.12	0.50	2.40	0.12	3.53	0.06

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
   "p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.41

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts with Low Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	0.50	0.50	0.94	0.50	0.10	0.50	0.14	0.50	1.43	0.23	0.45	0.50
PSSA Math	3.43	0.06	2.82	0.09	1.28	0.26	2.06	0.15	3.88	0.05	0.21	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	0.02	0.50	0.01	0.50	0.00	0.50	0.97	0.50	5.98	0.01	4.81	0.03
PSSA Math	3.74	0.05	3.62	0.05	0.58	0.50	3.54	0.06	2.56	0.11	0.75	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
   "p"is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.42

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts with High Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	0.00	0.50	0.09	0.50	0.51	0.50	1.00	0.50	0.81	0.50	0.02	0.50
PSSA Math	0.22	0.50	0.58	0.50	0.39	0.50	3.11	0.07	2.54	0.11	0.15	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	3.74	0.05	1.60	0.20	4.96	0.02	0.02	0.50	0.03	0.50	5.57	0.02
PSSA Math	1.85	0.17	1.98	0.16	0.04	0.50	0.61	0.50	7.29	0.01	0.48	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
   "p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.43

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts with Low Baseline Word Attack and Low Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	1.79	0.18	3.49	0.06	0.19	0.50	0.34	0.50	3.44	0.06	6.30	0.01
PSSA Math	5.68	0.02	5.72	0.02	1.47	0.22	1.93	0.16	4.68	0.03	1.44	0.23

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	0.00	0.50	0.06	0.50	0.57	0.50	0.25	0.50	5.03	0.02	11.34	0.00
PSSA Math	0.65	0.50	0.55	0.50	0.09	0.50	0.62	0.50	2.25	0.13	0.08	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
    "p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.44

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts With Low Baseline Word Attack and High Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	0.06	0.50	0.11	0.50	3.46	0.06	0.02	0.50	0.13	0.50	0.01	0.50
PSSA Math	1.24	0.27	2.50	0.11	0.78	0.50	0.70	0.50	1.83	0.17	7.98	0.00

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	2.40	0.12	1.02	0.31	1.84	0.17	0.14	0.50	0.12	0.50	1.18	0.28
PSSA Math	0.04	0.50	0.94	0.50	0.46	0.50	0.22	0.50	6.53	0.01	0.10	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
"p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.45

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts with High Baseline Word Attack and High Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*
PSSA Reading	0.02	0.50	0.00	0.50	0.13	0.50	9.23	0.00	7.43	0.01	0.02	0.50
PSSA Math	0.44	0.50	0.69	0.50	0.09	0.50	6.40	0.01	1.79	0.18	6.68	0.01

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*
PSSA Reading	5.03	0.02	3.44	0.06	1.20	0.27	0.63	0.50	0.21	0.50	4.80	0.03
PSSA Math	2.06	0.15	1.66	0.19	0.47	0.50	0.00	0.50	0.96	0.50	3.62	0.05

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
"*p*" is the *p*-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.46

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts Eligible for Free or Reduced Price School Lunch

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	7.20	0.01	2.53	0.11	23.11	0.00	2.45	0.11	0.31	0.50	2.54	0.11
PSSA Math	0.29	0.50	0.00	0.50	0.65	0.50	0.74	0.50	4.45	0.03	0.49	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	33.48	0.00	15.76	0.00	28.39	0.00	14.79	0.00	5.88	0.01	0.86	0.50
PSSA Math	2.25	0.13	2.23	0.13	1.22	0.27	3.09	0.07	1.51	0.22	0.71	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
  "p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.47

Test Statistics and P-values of Impacts for 3rd and 5th Grade Cohorts Not Eligible for Free or Reduced Price School Lunch

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	2.47	0.11	1.79	0.18	0.70	0.50	0.25	0.50	0.82	0.50	1.62	0.20
PSSA Math	1.05	0.31	2.10	0.14	0.97	0.50	0.18	0.50	3.62	0.05	0.07	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	1.40	0.23	1.15	0.28	0.24	0.50	0.00	0.50	0.62	0.50	0.73	0.50
PSSA Math	0.43	0.50	1.86	0.17	1.68	0.19	0.11	0.50	9.48	0.00	0.00	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
    "p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.48

Tests for Differences in Impacts for 3rd and 5th Grade Cohorts

	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	0.33	0.50	0.72	0.40	0.44	0.50	0.13	0.50	1.89	0.17	0.00	0.50
PSSA Math	5.97	0.01	6.41	0.01	1.33	0.25	0.02	0.50	16.17	0.00	1.38	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
   "p"is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.49

Tests for Differences in Impacts for 3rd and 5th Grade Cohorts with Low Baseline Word Attack Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	0.80	0.50	4.15	0.04	1.49	0.22	0.99	0.50	0.39	0.50	3.29	0.07
PSSA Math	6.08	0.01	7.35	0.01	0.00	0.50	0.30	0.50	2.01	0.15	15.30	0.00

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	0.09	0.50	0.32	0.50	0.31	0.50	0.00	0.50	0.39	0.50	2.28	0.13
PSSA Math	1.04	0.31	0.50	0.50	0.56	0.50	0.01	0.50	0.25	0.50	0.70	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
   "p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.50

Tests for Differences in Impacts for 3rd and 5th Grade Cohorts with Low Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	0.44	0.50	0.59	0.50	0.00	0.50	0.43	0.50	0.11	0.50	0.31	0.50
PSSA Math	2.44	0.11	1.56	0.21	1.51	0.22	5.34	0.02	0.29	0.50	0.14	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	1.85	0.17	0.84	0.50	1.53	0.21	0.26	0.50	0.50	0.50	13.53	0.00
PSSA Math	0.42	0.50	0.38	0.50	0.08	0.50	1.90	0.16	0.76	0.50	0.27	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
"p"is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.51

Tests for Differences in Impacts for 3rd and 5th Grade Cohorts with Low Baseline Word Attack and Low Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	2.66	0.10	3.99	0.04	0.04	0.50	1.27	0.26	1.17	0.28	7.10	0.01
PSSA Math	5.18	0.02	4.98	0.02	1.48	0.22	2.77	0.09	1.78	0.18	1.51	0.22

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	0.50	0.50	0.62	0.50	0.00	0.50	0.04	0.50	1.53	0.21	19.95	0.00
PSSA Math	0.01	0.50	0.02	0.50	0.00	0.50	0.22	0.50	0.01	0.50	0.37	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
   "p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.52

Tests for Differences in Impacts for 3rd and 5th Grade Cohorts With Low Baseline Word Attack and High Peabody Picture Vocabulary
Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	0.14	0.50	0.32	0.50	0.26	0.50	0.58	0.50	0.01	0.50	0.03	0.50
PSSA Math	0.03	0.50	1.41	0.23	1.78	0.18	2.14	0.14	0.23	0.50	7.56	0.01

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	0.03	0.50	0.10	0.50	0.17	0.50	0.21	0.50	0.13	0.50	0.08	0.50
PSSA Math	3.45	0.06	2.23	0.13	1.10	0.29	1.12	0.29	0.00	0.50	2.04	0.15

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
   "p"is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.53

Tests for Differences in Impacts for 3rd and 5th Grade Cohorts with High Baseline Word Attack and High Screening Peabody Picture Vocabulary Test Scores

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	0.40	0.50	0.01	0.50	2.00	0.15	1.41	0.23	3.21	0.07	0.02	0.50
PSSA Math	3.44	0.06	2.51	0.11	0.46	0.50	0.67	0.50	0.00	0.50	8.81	0.00

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	p	Test	p	Test	p	Test	p	Test	p	Test	p
PSSA Reading	1.33	0.25	1.07	0.30	0.12	0.50	0.01	0.50	0.01	0.50	4.91	0.03
PSSA Math	0.39	0.50	0.12	0.50	0.55	0.50	0.04	0.50	0.31	0.50	0.08	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
    "p" is the p-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

Table H.54

Tests for Differences in Impacts for 3rd and 5th Grade Cohorts Eligible for Free or Reduced Price School Lunch

Grade 3 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*
PSSA Reading	7.95	0.01	3.99	0.04	8.21	0.00	1.79	0.18	1.08	0.30	5.49	0.02
PSSA Math	2.04	0.15	1.75	0.18	0.09	0.50	0.21	0.50	5.63	0.02	0.60	0.50

Grade 5 Cohort	All Interventions ABCD		Word-level Interventions BCD		Failure Free Reading A		Spell Read B		Wilson Reading C		Corrective Reading D	
	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*	Test	*p*
PSSA Reading	5.88	0.01	2.66	0.10	3.81	0.05	7.91	0.01	0.34	0.50	0.07	0.50
PSSA Math	0.43	0.50	0.03	0.50	1.94	0.16	1.18	0.28	2.69	0.10	0.16	0.50

Note: "Test" refers to the Chi-Square test statistics for a general linear hypothesis test (Wald Test) with 1 degree of freedom.
"*p*" is the *p*-value produced by HLM6 software and is automatically set to 0.5 for values that exceed 0.5.

APPENDIX I

ESTIMATED R-SQUARED VALUES AND INTRACLASS CORRELATIONS

Table I.1

Estimated R-squared

| Test | R-squared | |
	Grade 3 Cohort	Grade 5 Cohort
Word Attack	0.45	0.39
TOWRE PDE	0.43	0.51
Word Identification	0.48	0.55
TOWRE SWE	0.55	0.47
AIMSweb	0.57	0.69
Passage Comprehension	0.39	0.24
GRADE	0.27	0.20

Note: The R-squared is the squared correlation between baseline and second follow-up test scores.

Table I.2

Estimated Intraclass Correlations

| | Intraclass Correlation | |
Test	Grade 3 Cohort	Grade 5 Cohort
Word Attack	0.17	0.24
TOWRE PDE	0.17	0.12
Word Identification	0.18	0.15
TOWRE SWE	0.21	0.10
AIMSweb	0.27	0.23
Passage Comprehension	0.19	0.18
GRADE	0.25	0.23

APPENDIX J

SCIENTIFIC ADVISORY BOARD

The following individuals served on the Scientific Board of Directors of the Haan Foundation for Children, which coordinated the selection of the four chosen interventions and the funding of the interventions and evaluation.

APPENDIX K

PSSA DATA COLLECTION FORM

SCHOOL NAME:_____

Pennsylvania System of School Assessment Scores (2003-04) for the Power4Kids Reading Initiative: Third-Grade Students

For each student listed in the Excel file sent to you, please provide the information specified in the column headings. (An example is provided in the table below.) We recommend extracting the data electronically from your files into the Excel file we provided. Alternatively, you may choose to manually enter the data into the Excel file or onto this form. Thank you.

MPRID	Name		Reading							Math		
	Last	First	Points Achieved				Scaled Score	Standard Error	Percentile	Scaled Score	Standard Error	Percentile
			Learning to Read Independently	Reading Critically in All Content Areas	Reading, Analyzing, and Interpreting Literature	Total Reading						
12345678	Smith	John	15	12	17	44	1561	96	91	1539	79	88

K-3

SCHOOL NAME:_____

Pennsylvania System of School Assessment Scores (2003-04) for the Power4Kids Reading Initiative: Fifth-Grade Students

For each student listed in the Excel file sent to you, please provide the information specified in the column headings. (An example is provided in the table below.) We recommend extracting the data electronically from your files into the Excel file we provided. Alternatively, you may choose to manually enter the data into the Excel file or onto this form. Thank you.

| MPRID | Name | | Reading | | | | | | | | | Math | | |
| | Last | First | Points Achieved | | | | | | Scaled Score | Standard Error | Percentile | Scaled Score | Standard Error | Percentile |
			Reading Independ-ently	Reading Critically	Analyzing/ Interpreting Literature	English Language Charac-teristics	Research	Total Reading						
98765432	Brown	Jane	17	13	26	11	13	80	1753	99	98	1684	60	95

APPENDIX L

SAMPLE TEST ITEMS

This appendix provides sample test items for the main tests used in the analysis.

Woodcock Reading Mastery Test-R, Word Attack Test (sample item)
Woodcock Reading Mastery Tests – Revised [NU] (WRMT-R/NU) Forms G and H. 1998. American
Guidance Services, Inc., Circle Pines, MN.

Tester says, "I want you to read some words that are not real words. I want you to tell me how they
sound."
Tester points to "tat" on the subject page.
Tester says, "How does that word sound?"

Woodcock Reading Mastery Test-R, Word Identification Test (sample item)
Woodcock Reading Mastery Tests – Revised [NU] (WRMT-R/NU) Forms G and H. 1998. American
Guidance Services, Inc., Circle Pines, MN.

Tester points to the word "is" on the subject page.
Tester says, "What is the word?"

Woodcock Reading Mastery Test-R, Passage Comprehension Test (sample item)
Woodcock Reading Mastery Tests – Revised [NU] (WRMT-R/NU) Forms G and H. 1998. American
Guidance Services, Inc., Circle Pines, MN.

Tester points to the sentence on the subject page and says, "This says, 'The cat is playing with
a…(pause)."

Sample A. The cat is playing with a _____.

Correct: ball
Incorrect: yarn

Tester points to the blank space in the sentence and says, "What word belongs in the blank space?

The cat is playing with a _____.

Test of Word Reading Efficiency (TOWRE), Sight Word Efficiency Test (practice items)
Test of Word Reading Efficiency (TOWRE), Forms A and B. 1999. PRO-ED, Inc, Austin, TX.

Tester says, "I want you to read some lists of words as fast as you can. Let's start with the practice
list. Begin at the top, and read down the list as fast as you can. If you come to a word you cannot read,
just skip it and go to the next word."
Practice Words: on, my, bee, old, warm, bone, most, spell

Test of Word Reading Efficiency (TOWRE), Phonemic Decoding Efficiency (practice items)
Test of Word Reading Efficiency (TOWRE), Forms A and B. 1999. PRO-ED, Inc, Austin, TX.

Tester says, "Now I want you to read some words that are not real words. Just tell me how they sound. I want you to read them as fast as you can. Let's start with the practice list. Begin at the top, and read down the list as fast as you can. If you come to a made-up word you cannot read, just skip it and go to the next word."
Practice Words: ba (b<u>a</u>t, f<u>a</u>te, pizz<u>a</u>), um (<u>u</u>mpire), fos (f<u>o</u>ssil), gan (g<u>a</u>nder), rup (r<u>u</u>pture), masp (cl<u>a</u>sp), luddy (m<u>u</u>ddy), dord (f<u>o</u>rd).

Group Reading Assessment and Diagnostic Evaluation Passage Comprehension (GRADE)
(sample item, third grade)
Group Reading Assessment and Diagnostic Evaluation (GRADE), Form A. 2001. American Guidance
    Services, Inc., Circle Pines, MN. Level 3 used for third grade; Level 5 used for fifth grade.

On hot days, Meg likes to jump in the lake. She stays in the cold water as long as she can. Meg's dad says, "Come on, little fish. It's time for lunch."

    1.   What does Meg do on hot days?
        a.   She goes for a bike ride.
        b.   She stays in the house.
        c.   She jumps in the lake.
        d.   She goes fishing.

AIMSWeb Oral Reading Passages
AIMSWeb Standard Reading Assessment Passages. 2001. Edformation, Inc. Eden Prairie, MN.

Third grade, Passage 2 (Copyright 2001 Edformation, Inc. All Rights Reserved)

Billy was sitting on the sidewalk curb holding his favorite old baseball glove.

"Hey, Billy!" he heard. "Weren't you supposed to meet me half an hour ago at the park? Why are you sitting here instead of moving?"

"I'm waiting," Billy replied.

"Waiting for what?" I asked.

"I'm waiting for Mr. Sanchez to leave for work. It shouldn't be much longer."

"Billy, your Mom said it was all right for you to play ball with me at the park. I don't understand why you're waiting for Mr. Sanchez."

Billy sighed as he explained, "Well, Mr. Sanchez's car is parked in the driveway, right across the street. See? And the park is across the street and down the block."

I shook my head because I didn't understand what Billy was talking about. This morning he was excited about playing baseball with the guys. He was a pretty good shortstop, even though he wasn't quite five years old.

"I know where the park is and so do you. So explain to me again why you are sitting here?"

"I already told you. I'm waiting for Mr. Sanchez," replied Billy. I looked across the street. There was no sign of Mr. Sanchez coming out of his house.

"Mom said I can't cross the street if I see any cars," Billy continued, "and I see Mr. Sanchez's car. It's right there in his driveway!"

"Oh, Billy!" I laughed. "I'm sure your mom meant you should not cross the street if you see any cars driving on the road! She just wants to make sure that a moving car doesn't hit you. She's not worried about the parked cars! Come on. You can walk with me to the park!"

"Oh, Sam. You're so smart. Thanks for being my friend. Let's go play ball."