Department of
Research and
Evaluation
AUSTIN INDEPENDENT SCHOOL DISTRICT

Research Brief

Background. The state of Texas mandates that districts with at least 15 or more eligible 4-year-old students provide a half-day prekindergarten (pre-K) program to prepare students developmentally for the rigors of kindergarten and beyond. In 2012–2013, Austin Independent School District (AISD) provided eligible students with full-day pre-K across 65 elementary campuses, the Lucy Read Prekindergarten Demonstration School, the Anita Uphaus Early Childhood Center, the Dobie Prekindergarten Center, and Webb Primary Center. This report evaluates students' achievement in the mandatory AISD pre-K program during the 2012–2013 school year, based on Peabody Picture Vocabulary Test-IV (PPVT) and Test de Vocabularío en Imagenes Peabody (TVIP) results.

Pre-literacy skills. The first 5 years of a child's life are influential with respect to his or her later school and life success (National Research Council, 2001). Decades of research indicate that early exposure to high-quality educational opportunities benefits young children by preparing them to enter school. Students' early language skills lay the foundation for later reading and writing in school, and students with poorer language skills are more likely to have difficulties learning to read than are students with good language skills (Whitehurst & Lonigan, 1998). Studies have shown that students who have difficulty learning to read in the earliest years of schooling tend to continue having difficulties over time (Scarborough, 2001; Snow, Burns, & Griffin, 1998). Vocabulary recognition is a precursor to the development of reading skills.

The PPVT and the TVIP measure knowledge of receptive vocabulary in English and Spanish, respectively (Dunn & Dunn, 2007). Receptive (i.e., hearing) vocabulary is one domain of language development. PPVT and TVIP pre- and post-assessments were administered by Department of Research and Evaluation (DRE) staff to pre-K students in the Lanier and Travis vertical teams, eight dual-enrollment classrooms at Dobie, and one randomly selected classroom at each of the remaining district campuses with a pre-K program. A total of 1,641[1] pre-K students were tested in both semesters. Of the pre-K students tested, 34% were English speakers in the state-mandated pre-K program (*n* = 552); 53% were Spanish-speaking English language learners (ELLs, *n* = 870); 2% were Spanish-

Program description. AISD provides a full-day pre-K program for all children who are 4-years-old on or before September 1st of the current school year and who meet one of the following eligibility criteria:

- Qualify for free or reduced-price lunch program
- English language learner
- Homeless
- Child of an active-duty military member or a military member who was injured or killed in service
- Reside or have ever resided in foster care

(TEC §29.153, [b])

The AISD pre-K program is designed to provide instructional activities to meet the individual needs of young children and stimulate learning in all areas of physical, social, emotional, and cognitive development.

[1] See technical note in Appendix A.

PPVT-IV and TVIP Interpretation

The standard scores for the PPVT and TVIP are based on age norms from samples of native speakers in English and Spanish, respectively. The tests are best interpreted in the native language of the student.

To make group comparisons and estimates of students' gains and losses, standardized scale scores were converted into normal curve equivalent (NCE) scores. NCE scores have a mean of 50 and a standard deviation of 21.06 points (i.e., the national normal distribution of scores). For the national samples on which the scores are based, the majority of students (68%) scored between 29 and 71 NCEs.

Interpretation of Normal Curve Equivalent (NCE) Scores

Range of performance	NCE scores	National percentile
Very low	1–4	1–11
Low	25–35	12–25
Low average	36–44	26–39
Average	45–55	40–60
High average	56–64	61–74
High	65–75	75–88
Very high	76–99	89–99

When interpreting NCE average gain (or loss), note that a zero shows that the students' growth did not differ from the national average expected growth; a zero does not indicate no growth. Because these tests are age normed, a student must have a raw score about 8 to 10 points higher in the spring to receive the same standard score as in the fall.

Deviations of ±4 NCE represent a significant growth (or loss), compared with national expected growth, and deviations of ±11 represent a very high significant gain (or loss).

speaking non-ELLs (n = 32) tested in Spanish; 8% were students who were tested in English, but whose home language is not English (n = 124); and 4% were tuition program students (n = 65).

AISD pre-K advances students to the national average in receptive vocabulary. Pre-K programs are designed to prepare students for kindergarten who might otherwise not have access to early educational programs. A quality pre-K program contributes to the reduction of the achievement gap in school performance that students face due to socioeconomic and language barriers. The goal of AISD pre-K is to raise students' academic performance to the average range of their peers. The assumption is that students who advance to at least the average range in the test of their language of instruction will be ready for future literacy learning in kindergarten.

PPVT and TVIP scores are best interpreted in the students' native language. For the PPVT, students are compared with a national sample of native English speakers. For the TVIP, students are compared with a national sample of native Spanish speakers.

In Fall 2012, AISD pre-K students in the mandatory program, on average, had a NCE score of 38.1 in their native language (Figure 1a). An estimated 69% of entering AISD pre-K students in the mandatory program (rather than an expected 50%) had scores below the national mean score; 40% were one standard deviation below average. The median equivalent age (3 years, 11 months) of mandatory pre-K students' native language receptive vocabulary ability was 8 months below the median actual age (4 years, 7 months).

In Spring 2013, pre-K students' distribution of PPVT and TVIP scores more closely approximated the national distribution than it did in Fall 2012 (Figure 1b). The mean NCE was 51.7 (i.e., above the national average), and 13% of pre-K students were one standard deviation below average. The median age-equivalent score was 3 months above the median actual age (5 years, 2 months).
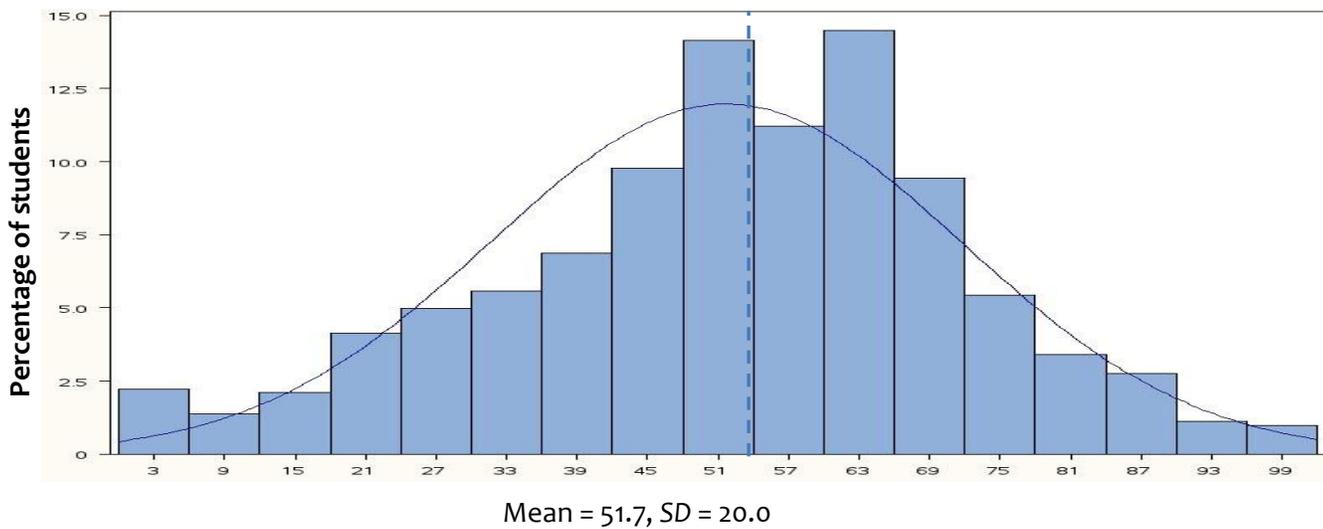
Figure 1. Mandatory Prekindergarten (pre-K) Students' Peabody Picture Vocabulary Test (PPVT-IV) and Test de Vocabulario en Imagenes Peabody (TVIP) Performance Scores in Native Language, by Semester, 2012–2013

Normal curve equivalent (NCE) scores

a. **Fall 2012**
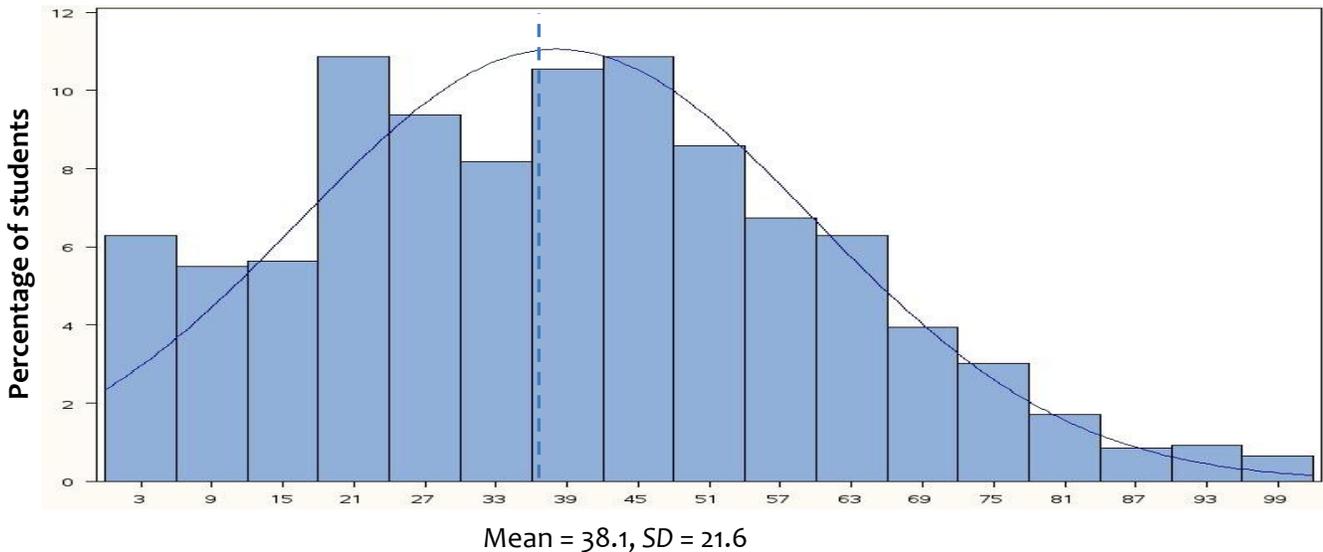
Mean = 38.1, *SD* = 21.6

b. **Spring 2013**

Mean = 51.7, *SD* = 20.0

Source. AISD student records, Department of Research and Evaluation

Note. Results for the TVIP were based on a weighted sample (see Appendix A). Original sample, *n* = 1,452. The dotted line represents the national mean NCE score.

PPVT and TVIP, according to ELL status. The average PPVT-IV NCE score for English-speaking students in Fall 2012 was 40.2, while the average TVIP NCE score for Spanish-speaking ELLs was 36.7. In Spring 2013, the average NCE score for English-speaking pre-K students was 47.7, with an average gain of 7.4 points, indicating greater than expected growth. For Spanish-speaking ELLs, the posttest scores averaged 53.7, with an average gain of 17 points.

In Spring 2013, more than 50% of students scored above the national average on the PPVT and TVIP.

Approximately 84% of students were expected to be in the average range or higher in the standard normal distribution in which the scores were based. However, an estimated 70% of English-speaking students and 59% of Spanish-speaking ELLs scored in the average range or higher in Fall 2012. In Spring 2013, 86% to 87% of pre-K students in the state-mandated program scored in the average range or higher in their native language (Figure 2).

Sixty-one percent of English-speaking students and 72% of Spanish-speaking ELLs who scored below the average range in the fall advanced to the average range or higher on the PPVT and TVIP, respectively, in the spring.

Figure 2. Mandatory Prekindergarten Students' Peabody Picture Vocabulary Test (PPVT-IV) and Test de Vocabularío en Imagenes Peabody (TVIP) Performance, by Native Language and Semester, 2011–2012
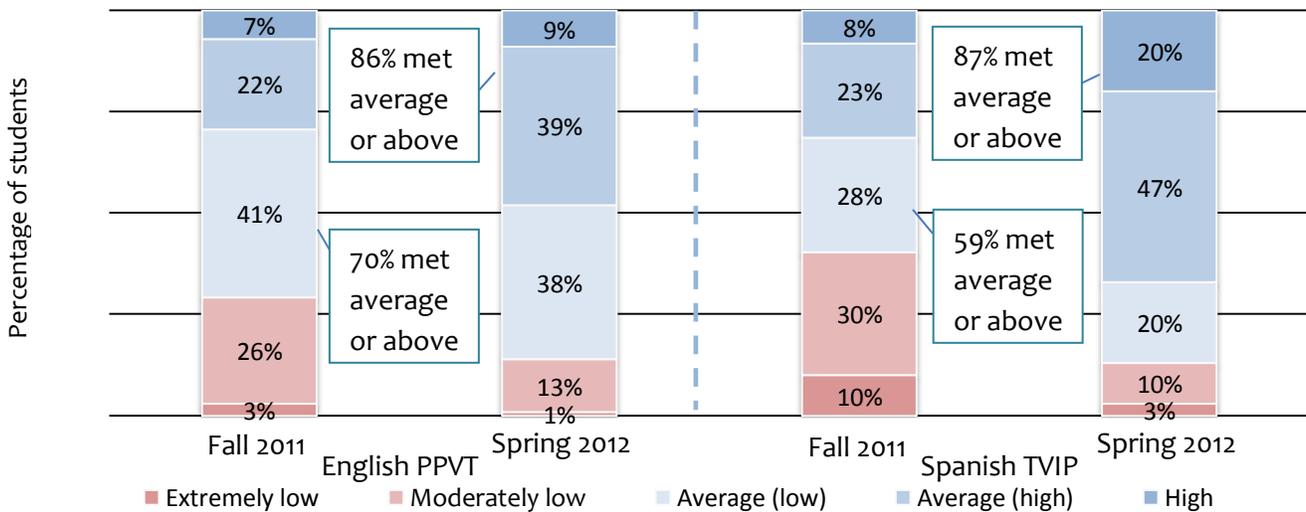
Source. AISD student records, Department of Research and Evaluation

Note. Results for the TVIP were based on a weighted sample (see Appendix A). Original sample: English, *n* = 552; *Spanish, n* = 870.

In general, the AISD mandatory pre-K program accelerated pre-K students' growth in receptive vocabulary above expected growth. On average, Spanish-speaking students gained an estimated 16 months of growth in equivalent age in a 7-month period. An estimated 58% of Spanish-speaking ELLs improved their age-equivalency score by more than 7 months (i.e., faster than the expected growth rate) in receptive vocabulary.

In a 7-month period, tested pre-K students gained the equivalent of 11 to 16 months of growth in receptive vocabulary, on average.

English-speaking students gained 11 months of age-equivalent growth, on average. Approximately 33% of English-speaking students improved their age-equivalency score by more than 7 months.

Growth Scale Values (GSVs) for non-native English speakers. GSVs measure students' knowledge based on an absolute scale rather than a normed scale. This means that as a student's vocabulary grows, the GSV will increase. Unlike raw scores, GSVs can be used across different versions of assessment forms, age groups, and grade levels to measure growth. In general, an 8-point increase in the GSV scale for students under the age of 12 is considered significant (i.e., statistically different from the previous time-period; Dunn & Dunn, 2007). Unlike NCEs, GSVs cannot inform about expected growth. (Expected growth is determined by a reference group.) The average non-native English speaker's growth in English vocabulary was 19 points (Figure 3). Eighty-one percent (*n* = 101) of the non-native English speakers tested with the PPVT demonstrated an 8-point gain or more.

Figure 3. Growth Scale Values for Non-Native English Speakers, by Semester (*n* = 124)
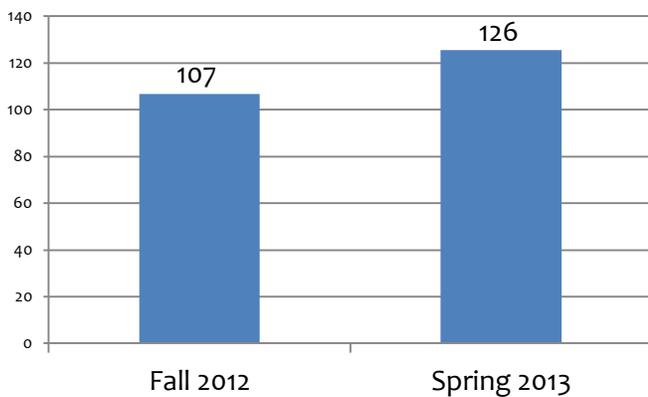
Source. AISD student records

Findings suggest pre-K students benefited from multiple program interventions (e.g., one-way dual language and the Texas Literacy Initiative) and early childhood centers with Head Start services.

Peabody results for tuition pre-K program students. Sampled tuition pre-K students' (*n* = 65) average NCE was 69.4 in Fall 2012, nearly one standard deviation above the national average. In Spring 2013, the average NCE for these students was 70.6, indicating expected growth. Using the GSV measure, 51% of tuition students (*n* = 33) made significant growth (i.e., at least an 8-point gain in the GSV scale). However, the average GSV gain for the sample was 7.2 points.

Early childhood centers. Three new early childhood centers were opened in AISD in Fall 2012. As part of the Texas Literacy Initiative (TLI), all pre-K students at Read and Uphaus were tested. Eight classrooms at Dobie with pre-K students dually enrolled with Head Start also were tested. Results are provided in Table 1. Webb was not included due to sample size criteria. On average, pre-K students at Dobie and Read scored lower in the fall on the Peabody than did students who were not enrolled in a childhood center, indicating these schools had greater student academic needs, in general, than did other pre-K programs across the district. Read and Dobie students, on average, had greater gains per student than did students who were not enrolled at a childhood center. Pre-K students at Uphaus had higher scores on the Peabody in Fall 2012 than did students district wide. In Spring 2013, Uphaus students scored in the average to high range.

Table 1. Peabody Normal Curve Equivalent (NCE) Results for State-Mandated Prekindergarten Students, by English Language Learner (ELL) Status and Early Childhood Center, 2012–2013

	# students tested both	Avg. NCE Fall	Avg. NCE Spring	Average gain per student
District wide	**1,526**	**38.2**	**51.7**	**13.5**
Dobie*	113	29.5	45.5	15.9
Read	343	30.1	49.9	19.1
Uphaus	66	44.7	56.8	12.2
Non-center	991	41.4	52.8	11.4

Source. AISD student records, Department of Research and Evaluation

* Only Dobie pre-K classrooms dually enrolled with Child, Inc. (i.e., Head Start) were included. Webb Primary Center was not included because only one classroom was tested. Webb was included in the district-wide calculations. Presented results were based on weighted scores (see Appendix).

Multiple program effect. To determine classroom-level effects, school-level effects, and tester bias, DRE staff used hierarchical linear modeling (HLM). After controlling for tester bias (see Appendix A), classroom-level differences accounted for 11% of the total variation in pre-K student performance (i.e., student performance differed in each classroom). DRE staff used HLM to determine the effect of various programs on students' receptive vocabulary growth. Programs considered in the analysis included one-way dual language; two-way dual language; dual-enrollment classrooms with Child, Inc. (i.e., Head Start); early education centers (i.e., Read, Dobie, Webb, and Uphaus); and TLI classrooms (i.e., Travis and Lanier vertical teams). Students' ELL status and special education status were included in the models.[2] In addition to examining each program's effect on student performance, DRE staff also looked at cross-program effects for two programs at a time.[3] Table 2 provides the estimated program effect in NCEs for each program.

---

[2] Economic disadvantage also was considered; however, because all the students in some programs were economically disadvantaged, this factor was removed from the models to ensure convergence. In the student-level effects only model, economic status was not significant.

[3] Some classrooms participated in three or four programs. Interaction effects for three or more programs were not examined.

Table 2. Peabody Normal Curve Equivalent (NCE) Program Effect Estimates, by Program, 2012–2013

	ECC	Head Start	One-way	Two-way	TLI
Early childhood center (ECC)	.58 (.64)	-	-	-	-
Head Start dual enrollment	7.55 (.06)	1.60 (.29)	-	-	-
One-way dual language (DL)	1.30 (.62)	-0.38 (.92)	5.61 (<.01)	-	-
Two-way DL	NA	NA	NA	-3.16 (.16)	-
Texas Literacy Initiative (TLI)	3.53 (.21)	5.27 (.10)	4.66 (.03)	.05 (.99)	1.51 (.15)

Source. AISD student Peabody records

Note. Estimates are in terms of NCE points. Models were run separately. Each model controlled for fall NCE, English language learner status, special education status, tester bias, and teacher effect. Cells highlighted in blue are significant. Cells highlighted in green are close to statistical significance. Probabilities are noted in parentheses. NA is not applicable (i.e., no classrooms were represented).

In general, the one-way dual language program increased pre-K students' spring NCE score by 5.6 points. Notably, the Peabody assessment measures receptive vocabulary in students' native language, and results do not represent students' growth in their second language (i.e., English). Also, one-way classrooms in the Lanier and Travis vertical teams (i.e., TLI schools) increased pre-K students' NCE scores by 4.7 points, on average (Table 2). Table 3 provides average Peabody scores for mandatory pre-K students in TLI vertical teams. Students in the Lanier vertical teams had greater growth in NCEs than did students in non-TLI schools. In general, pre-K students in the Travis vertical team had higher Peabody scores than pre-K students did in the rest of the district.

Table3. Peabody Normal Curve Equivalent (NCE) Results for State-Mandated Prekindergarten Students, by Texas Literacy Initiative (TLI) Vertical Team, 2012–2013

	# students tested both	Avg. NCE Fall	Avg. NCE Spring	Average gain per student
District wide	**1,526**	**38.2**	**51.7**	**13.5**
Lanier vertical team	401	31.5	50.8	19.3
Travis vertical team	412	44.0	54.7	10.6
Non-TLI vertical teams	713	38.6	50.4	11.8

Source. AISD student records, Department of Research and Evaluation

Child, Inc., collaborative classrooms. Although the cross-program effect was slightly above the probability threshold of .05 (i.e., .0578), this may reflect some students receiving services from both Head Start (i.e., Child, Inc.) and AISD pre-K at an early childhood center. On average, dually enrolled pre-K students scored 7.6 points higher on the spring Peabody assessment than did other tested students. The AISD and Child, Inc., collaboration provides both Head Start and pre-K services to families who qualify for both services. AISD

hosted 22 collaborative classrooms. Each classroom received all of the standard pre-K services provided in AISD as well as Head Start services, including mental and medical benefits, family education services, funded field trips, classroom supplies, and inclusion within a classroom of no more than 20 students. Head Start collaborative classrooms have an assigned AISD teacher and teacher assistant. The collaboration seeks to ensure kindergarten readiness for all students by providing additional services to students who are most in need. The findings suggest that the combination of AISD pre-K at the centers and Head Start programs worked together to increase students' test performance growth faster than the expected national average.

Conclusion. The AISD mandatory pre-K program continued to achieve its goal of raising the majority of pre-K students to grade-level expectations by the end of the year, preparing them for success in kindergarten. On average, the pre-K program helped students progress an average of 14 months in age-equivalent ability in receptive vocabulary. At the end of the year, an estimated 10% of pre-K students scored below the average range, a decrease from 40% in the fall. (In a normal distribution, 16% of students are expected to be below the average range.) Overall, the data show that the AISD pre-K program combined with other services (i.e., dual language, Head Start, early childhood center model) was effective in increasing growth in receptive vocabulary. AISD should continue to support funding of a full-day program for students to ensure that students continue to enter kindergarten ready to learn on grade level.

References.

Dunn, L., & Dunn, L. (2007). *Examiner's manual for the Peabody picture vocabulary test* (4th ed.)*.* San Antonio, TX: Pearson PsychCorp.

National Research Council, Bowman, B., Donovan, M. S., & Burns, S. (Eds.). (2001.) *Eager to learn: Educating our preschoolers.* Washington, DC: National Academy Press.

Scarborough, H. S. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. B. Neuman, & D. K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 97–110). New York, NY: Guilford Press.

Snow, C. E., Burns, M. S., & Griffin, P. (1998)*. Preventing reading difficulties in young children*. Washington, DC: National Research Council, National Academy.

Texas Education Code, §29.153, (b). (2007).

Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development,* 69, 848–872.

Appendix A

Technical Note

Originally, 1,825 pre-K students were tested in both the fall and spring semesters. Tester bias was detected in the spring data using HLM. The error was caused by missing basals (i.e., lower score threshold) on the Spanish assessment (i.e., TVIP). Invalid test scores (n = 182) were removed from the sample. Missing basals

tend to inflate test scores because items below the basal are assumed to be correct. Lower-performing students were more likely to be removed from the sample because they were more likely to have missed previous items assumed correct but not tested. DRE staff weighted TVIP scores for each economic status, vertical team, ELL status, and fall scale scores to statistically adjust for the systematic error. By weighting scores based on fall scale scores, lower-performing students were provided greater weight to adjust for any missing data. Overall, the average district TVIP score decreased by 3 points using this adjustment.